PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/574, C12Q 1/68

A2

(11) International Publication Number: WO 98/53319

(43) International Publication Date: 26 November 1998 (26.11.98)

(21) International Application Number: PCT/US98/10277

(22) International Filing Date: 20 May 1998 (20.05.98)

(30) Priority Data:
60/047,352 21 May 1997 (21.05.97) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 60/047,352 (CON)
Filed on 21 May 1997 (21.05.97)

(71) Applicant (*for all designated States except US*): THE JOHNS HOPKINS UNIVERSITY [US/US]; Suite 2–100, 2024 E. Monument Street, Baltimore, MD 21205 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): VOGELSTEIN, Bert [US/US]; The Johns Hopkins University, Suite 2–100, 2024 E. Monument Street, Baltimore, MD 21205 (US). KINZLER, Kenneth, W. [US/US]; The Johns Hopkins University, Suite 2–100, 2024 E. Monument Street, Baltimore, MD 21205 (US).

(74) Agents: KAGAN, Sarah, A. et al.; Banner & Witcoff, Ltd., 11th floor, 1001 G Street, N.W., Washington, DC 20001–4597 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GENE EXPRESSION PROFILES IN NORMAL AND CANCER CELLS

(57) Abstract

As a step towards understanding the complex differences between normal and cancer cells, gene expression patterns were examined in gastrointestinal tumors. More than 300,000 transcripts derived from at least 45,000 different genes were analyzed. Although extensive similarity was noted between the expression profiles, more than 500 transcripts that were expressed at significantly different levels in normal and neoplastic cells were identified. These data provide insights into the extent of expression differences underlying malignancy and reveal genes that are useful as diagnostic or prognostic markers.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Gene Expression Profiles in Normal and Cancer Cells

This invention was made with support from the National Institutes of Health, Grant No. GM07309, CA57345, and CA62924. The U.S. government therefore retains certain rights in the invention.

TECHNICAL FIELD OF THE INVENTION

This invention is related to the diagnosis of cancer, and tools for carrying out such diagnosis.

BACKGROUND OF THE INVENTION

Much of cancer research over the past 50 years has been devoted to the analyses of genes that are expressed differently in tumor cells compared to their normal counterparts. Although hundreds of studies have pointed out differences in the expression of one or a few genes, no comprehensive study of gene expression in the cancer cell has been reported. It is therefore not known how many genes are expressed differentially in tumor versus normal cells, whether the bulk of these differences are cell autonomous rather than being dependent on the tumor microenvironment, and whether most differences are cell-type specific or tumor specific. Thus there is a need in the art for information on the molecular changes that occur in cells during cancer development and progression.

SUMMARY OF THE INVENTION

According to one embodiment of the invention, a method is provided for diagnosing colon cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 3;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be lower in the first sample than in the second sample.

According to another embodiment of the invention, another method is provided for diagnosing colon cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

In another embodiment of the invention an isolated and purified human nucleic acid molecule is provided. The molecule comprises a SAGE tag selected from SEQ ID NO:1-732.

In yet another aspect of the invention an isolated nucleotide probe is provided. The probe comprises at least 12 nucleotides of a human nucleic acid molecule, wherein the human nucleic acid molecule comprises a SAGE tag selected from SEQ ID NO: 1-732.

According to another aspect of the invention a method is provided for diagnosing pancreatic cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a pancreatic tissue suspected of being neoplastic and the second sample is of a normal human colon tissue, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

According to still another embodiment of the invention a method of diagnosing cancer in a sample suspected of being neoplastic is provided. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a tissue suspected of being neoplastic and the second sample is of a normal human tissue, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 5;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

According to another embodiment of the invention a method is provided to aid in the determination of a prognosis for a colon cancer patient. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic colonic tissue and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 3;

determining a poorer prognosis if the level of the at least one transcript is found to be lower in the first sample than in the second sample.

According to another aspect of the invention a method to aid in determining a prognosis for a patient with colon cancer is provided. The method comprises the steps of:

comparing the level of at least one transcript in a first tissue sample to a second sample, wherein the first sample is of a colonic cancer tissue and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

In yet another embodiment of the invention a method is provided for diagnosing colon cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 3;

identifying the first sample as neoplastic when the level of expression of the protein is found to be lower in the first sample than in the second sample.

In another aspect of the invention a method of diagnosing colon cancer in a sample suspected of being neoplastic is provided. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript

identified by a tag selected from the group consisting of those shown in Table 2;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

According to another embodiment of the invention a method is provided to aid in determining a prognosis of a patient having pancreatic cancer. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic pancreatic tissue and the second sample is of a normal human colon tissue, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4;

determining a poorer prognosis if transcription is found to be higher in the first sample than in the second sample.

In yet another aspect of the invention a method to aid in providing a prognosis for a cancer patient is provided. The method comprises the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic tissue and the second sample is of a normal human tissue of the same tissue type, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 5;

determining a poorer prognosis if transcription is found to be higher in the first sample than in the second sample.

According to still another aspect of the invention, a method is provided for diagnosing pancreatic cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of expression of at least one protein encoded by a transcript in a first sample of a tissue to a second sample, wherein the first sample is of a pancreatic tissue suspected of being neoplastic and the second sample is of a normal human colon tissue, wherein said protein is

encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

According to yet another aspect of the invention a method is provided for diagnosing cancer in a sample suspected of being neoplastic. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a tissue suspected of being neoplastic and the second sample is of a normal human tissue, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 5;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

In still another embodiment of the invention a method is provided to aid in the determination of a prognosis of a colon cancer patient. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic colonic tissue and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 3;

determining a poorer prognosis if the level of expression is found to be lower in the first sample than in the second sample.

In still another embodiment of the invention a method is provided to aid in determining a prognosis for a patient with colon cancer. The method comprises the steps of:

comparing the level of expression of at least one protein in a first tissue sample to a second sample, wherein the first sample is of a colonic cancer tissue and the second sample is of a normal human colonic tissue, and

wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 2;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

In still another aspect of the invention a method is provided to aid in determining a prognosis of a patient having pancreatic cancer. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic pancreatic tissue and the second sample is of a normal human colon tissue, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

According to even a further aspect of the invention a method is provided to aid in providing a prognosis for a cancer patient. The method comprises the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic tissue and the second sample is of a normal human tissue of the same tissue type, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 5;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

In still another embodiment of the invention a method of treating a cancer cell is provided. The method comprises the step of:

administering to a cancer cell an antibody which specifically binds to a protein encoded by a transcript identified by a tag selected from the group consisting of those shown in Tables 2, 4, and 5, wherein the antibody is linked to a cytotoxic agent.

In another aspect of the invention an antibody linked to a cytotoxic agent is provided. The antibody specifically binds to a protein encoded by a transcript identified by a tag selected from the group consisting of those shown in Tables 2, 4, and 5.

According to another aspect of the invention, a method of detecting colon cancer in a patient is provided. The method comprises the steps of:

comparing the level of at least one protein or transcript in a first body sample to a second body sample, wherein the first sample is a body sample of the patient and the second sample is of a normal human, wherein the protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

In another aspect of the invention a method of detecting pancreatic cancer in a patient is provided. The method comprises the steps of:

comparing the level of at least one protein or transcript encoded by a transcript in a first sample of a tissue to a second sample, wherein the first sample is of the patient and the second sample is of a normal human, wherein said protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown Table 4, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

Also provided by the present invention is a method of detecting cancer in a patient. The method comprises the steps of:

comparing the level of at least one protein or transcript in a first sample to a second sample, wherein the first sample is of patient and the second sample is of a normal human, wherein said protein is encoded by a

transcript and the transcript is identified by a tag selected from the group consisting of those shown Table 5, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

Additionally provided by the present invention is a method to aid in the determination of a prognosis for a colon cancer patient. The method comprises the steps of:

comparing the level of at least one protein or transcript in a first sample to a second sample, wherein the first sample is of a colon cancer patient and the second sample is of a normal human, wherein the protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown in Table 3, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein or transcript is found to be lower in the first sample than in the second sample.

Provided by another embodiment of the invention is a method to aid in determining a prognosis for a patient with colon cancer. The method comprises the steps of:

comparing the level of at least one protein or transcript in a first sample to a second sample, wherein the first sample is of a colonic cancer patient and the second sample is of a normal human, wherein the protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

According to still another aspect of the invention, a method to aid in determining a prognosis of a patient having pancreatic cancer is provided. The method comprises the steps of:

comparing the level of at least one protein or transcript in a first sample to a second sample, wherein the first sample is of a pancreatic cancer patient and the second sample is of a normal human, wherein said protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown Table 4, wherein said first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

Also provided by the present invention is a method to aid in providing a prognosis for a cancer patient. The method comprises the steps of:

comparing the level of expression of at least one protein or transcript in a first sample to a second sample, wherein the first sample is of a cancer patient and the second sample is of a normal human, wherein said protein is encoded by a transcript and the transcript is identified by a tag selected from the group consisting of those shown Table 5, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein or transcript is found to be higher in the first sample than in the second sample.

The present invention further includes antisense oligonucleotides complementary in whole or in part to SEQ ID NOS:1-732.

This invention also provides a method for screening for candidate agents that modulate the expression of a polynuleotide selected from the group consisting of the polynucleotides in SEQ ID NOS.1-732 or their respective complements, by contacting a test agent with a pancreatic or colon cell and monitoring expression of the polynucleotide, wherein the test agent which modifies the expression of the polynucleotide is a candidate agent.

The present invention provides the art with new methods and reagents for diagnosing and prognosing cancers. In addition, some of the newly disclosed genes may play an important role in the development of cancers.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. Comparison of expression patterns in colorectal cancers and normal colon epithelium. (FIG. 1A) A semi-logarithmic plot reveals 51 tags that were decreased more than 10 fold in primary CR cancer cells whereas 32 tags were increased more than 10 fold. 62,168 and 60,878 tags derived from normal colon epithelium and primary CR cancers, respectively, were used for this analysis. The relative expression of each transcript was determined by dividing the number of tags observed in tumor and normal tissue as indicated. To avoid division by 0, a tag value of 1 was used for any tag that was not detectable in one of the samples. These ratios were then rounded to the nearest integer and their distribution plotted on the abscissa. The number of genes displaying each ratio was plotted on the ordinate. Tu: CR tumors; NC: Normal colon. (FIG. 1B and FIG. 1C) Differentially expressed genes in colorectal cancers. The number of transcripts found to be differentially expressed ($P < 0.01$) are presented as Venn diagrams. Diagrams of transcripts that were decreased (FIG. 1B) or increased (FIG. 1C) in CR cancers compared to normal colon epithelium. Comparisons were between primary tumors and cells in culture as indicated.

Fig. 2. Northern blot analysis of genes differentially expressed in gastrointestinal neoplasia. Northern blot analysis was performed on total RNA (5 μg isolated from primary CR carcinomas (T) and matching normal colon epithelium (N), or pancreatic carcinomas. The top panel in each case show an

example of the ethidium bromide stained gels prior to transfer. The number of SAGE tags observed in the original analysis is indicated to the right of each blot. (FIG. 2A) Examples of transcripts that were decreased or increased in CR cancers. (FIG.2B) Examples of transcripts increased in pancreatic cancers (*10*). (FIG.2C) Examples of transcripts elevated in cancer which were or were not cancer type specific. Probes used for Northern blot analysis were as follows (Human SAGE Tag unique identifier, gene name, (GenBank accession number)): (FIG. 2A) H204104, Guanylin (M95714); H259108, (see Table 2); H1000193, (see Table 2); H998030, (see Table 2). (FIG. 2B) H294155, RIG-E (U42376); H560056, TIMP-1 (S68252). (FIG. 2C) H802810, EST338411 (W52120); H85882, 1-8D (X57351); H618841, GA733-1 (X13425).

Tables 2-5. Transcripts Differentially Expressed in Human Cancer.

Tag sequence represents the NlaIII site plus the adjacent 11 bp SAGE tag. Tag number indicates a SAGE UID (unique identifier). NC, TU, CL, PT, PC, refers to the number of the indicated tag observed in RNA isolated from normal colorectal epithelium, primary colorectal cancers, colorectal cancer cell lines, primary pancreatic cancers, or pancreatic cancer cell lines, respectively. The Accession and Gene Name refer to representative GenBank entries that contain the tag sequence.

Table 2 Transcripts increased in colorectal cancer.

Table 3 Transcripts decreased in colorectal cancer.

Table 4 Transcripts increased in pancreatic cancer.

Table 5 Transcripts increased in pancreatic and colorectal cancer.

DETAILED DESCRIPTION

The inventors have discovered sets of human genes which are either upregulated or downregulated in cancer cells, as compared to normal cells. Specifically, certain genes have been found to be upregulated or downregulated in colorectal and/or pancreatic cancer cells, when compared to normal colon

cells. These sets of differentially regulated genes can be used as diagnostic markers, either individually or in sets of, for example, 2, 5, 10, 20, or 30.

Genes whose expression was detected to be increased in colorectal cancer are shown in Table 2. Genes whose expression was detected to be decreased in colorectal cancer are shown in Table 3. Genes whose expression was detected as increased in pancreatic cancer are shown in Table 4. Genes whose expression was detected as increased in both pancreatic cancer and colorectal cancer are shown in Table 5. These latter genes likely play a role in neoplastic development generally.

Tag sequences, as provided herein, uniquely identify genes. This is due to their length, and their specific location (3') in a gene from which they are drawn. The full length genes can be identified by matching the tag to a gene data base member, or by using the tag sequences as probes to physically isolate previously unidentified genes from cDNA libraries. The methods by which genes are isolated from libraries using DNA probes are well known in the art. See, for example, Veculescu et al., Science 270: 484 (1995), and Sambrook et al. (1989), MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed. (Cold Spring Harbor Press, Cold Spring Harbor, New York). Once a gene or transcript has been identified, either by matching to a data base entry, or by physically hybridizing to a cDNA molecule, the position of the hybridizing or matching region in the transcript can be determined. If the tag sequence is not in the 3' end, immediately adjacent to the restriction enzyme used to generate the SAGE tags, then a spurious match may have been made. Confirmation of the identity of a SAGE tag can be made by comparing transcription levels of the tag to that of the identified gene in certain cell types.

In addition to the sequences shown in SEQ ID NOS: 1-732, or their complements, this invention also provides the anti-sense polynucleotide stand, e.g. antisense RNA to these sequences or their complements. One can obtain an antisense RNA using the sequences provided in SEQ ID NOS: 1-732 and the methodology described in Vander Krol et al. (1988) BioTechniques 6:958.

The invention also encompasses polynucleotides which differ from that of the polynucleotides described above, but which produce the same phenotypic effect, such as the allele. These altered, but phenotypically equivalent polynucleotides are referred to "equivalent nucleic acids." This invention also encompasses polynucleotides characterized by changes in non-coding regions that do not alter the phenotype of the polypeptide produced therefrom when compared to the polynucleotide herein. This invention further encompasses polynucleotides, which hybridize to the polynucleotides of the subject invention under conditions of moderate or high stringency.

The polynucleotides can be conjugated to a detectable marker, e.g., an enzymatic label or a radioisotope for detection of nucleic acid and/or expression of the gene in a cell. A wide variety of appropriate detectable markers are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred embodiments, one will likely desire to employ a fluorescent label or an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples. Briefly, this invention further provides a method for detecting a single-stranded polynucleotide identified by SEQ ID NOS.1-732 or its complement, by contacting target single-stranded polynucleotides with a labeled, single-stranded polynucleotide (a probe) which is at least 10 nucleotides of the complement of SEQ ID NOS: 1-732 (or the corresponding complement) under conditions permitting hybridization (preferably moderately stringent hybridization conditions) of complementary single-stranded polynucleotides, or more preferably, under highly stringent hybridization conditions. Hybridized polynucleotide pairs are separated from un-hybridized, single-stranded polynucleotides. The hybridized polynucleotide

pairs are detected using methods well known to those of skill in the art and set forth, for example, in Sambrook et al. (1989) supra.

The polynucleotides of this invention can be isolated using the technique described in the experimental section or replicated using PCR. The PCR technology is the subject matter of United States Patent Nos.4,683,195, 4,800,159, 4,754,065, and 4,683,202 and described in PCR: The Polymerase Chain Reaction (Mullis et al. eds, Birkhauser Press, Boston (1994)) or MacPherson et al. (1991) and (1994), supra, and references cited therein. Alternatively, one of skill in the art can use the sequences provided herein and a commercial DNA synthesizer to replicate the DNA. Accordingly, this invention also provides a process for obtaining the polynucleotides of this invention by providing the linear sequence of the polynucleotide, nucleotides, appropriate primer molecules, chemicals such as enzymes and instructions for their replication and chemically replicating or linking the nucleotides in the proper orientation to obtain the polynucleotides. In a separate embodiment, these polynucleotides are further isolated. Still further, one of skill in the art can insert the polynucleotide into a suitable replication vector and insert the vector into a suitable host cell (procaryotic or eucaryotic) for replication and amplification. The DNA so amplified can be isolated from the cell by methods well known to those of skill in the art. A process for obtaining polynucleotides by this method is further provided herein as well as the polynucleotides so obtained.

RNA can be obtained by first inserting a DNA polynucleotide into a suitable host cell. The DNA can be inserted by any appropriate method, e.g., by the use of an appropriate gene delivery vector or by electroporation. When the cell replicates and the DNA is transcribed into RNA; the RNA can then be isolated using methods well known to those of skill in the art, for example, as set forth in Sambrook et al. (1989) supra. For instance, mRNA can be isolated using various lytic enzymes or chemical solutions according to the procedures set forth in Sambrook et al. (1989), supra or extracted by nucleic-acid-binding resins following the accompanying instructions provided by manufactures.

Polynucleotides having at least 10 nucleotides and exhibiting sequence complementarity or homology to SEQ ID NOS: 1-732 find utility as hybridization probes. In some aspects, the full coding sequence of the transcript, i.e., for SEQ ID NOS: 1-732, are known. Accordingly, any portion of the known sequences available in GenBank, or homologous sequences, can be used in the methods of this invention.

It is known in the art that a "perfectly matched" probe is not needed for a specific hybridization. Minor changes in probe sequence achieved by substitution, deletion or insertion of a small number of bases do not affect the hybridization specificity. In general, as much as 20% base-pair mismatch (when optimally aligned) can be tolerated. Preferably, a probe useful for detecting the aforementioned mRNA is at least about 80% identical to the homologous region of comparable size contained in the previously identified sequences identified by SEQ ID NOS:1-732, which correspond to previously characterized genes or SEQ ID NOS:1-732, which correspond to known ESTs. More preferably, the probe is 85% identical to the corresponding gene sequence after alignment of the homologous region; even more preferably, it exhibits 90% identity.

These probes can be used in radioassays (e.g. Southern and Northern blot analysis) to detect, prognose, diagnose or monitor various pancreatic or colon cells or tissue containing these cells. The probes also can be attached to a solid support or an array such as a chip for use in high throughput screening assays for the detection of expression of the gene corresponding to one or more polynucleotide(s) of this invention. Accordingly, this invention also provides at least one of the transcripts identified as SEQ ID NOS:1-732, or its complement, attached to a solid support for use in high throughput screens.

The total size of fragment, as well as the size of the complementary stretches, will depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the complementary region may be varied,

such as between about 10 and about 100 nucleotides, or even full length according to the complementary sequences one wishes to detect.

Nucleotide probes having complementary sequences over stretches greater than 10 nucleotides in length are generally preferred, so as to increase stability and selectivity of the hybrid, and thereby improving the specificity of particular hybrid molecules obtained. More preferably, one can design polynucleotides having gene-complementary stretches of more than 50 nucleotides in length, or even longer where desired. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR technology with two priming oligonucleotides as described in U.S. Pat. No. 4,603,102 or by introducing selected sequences into recombinant vectors for recombinant production. A preferred probe is about 50-75 or more preferably, 50-100, nucleotides in length.

The polynucleotides of the present invention can serve as primers for the detection of genes or gene transcripts that are expressed in pancreatic or colon cells. In this context, amplification means any method employing a primer-dependent polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA-polymerases such as T7 DNA polymerase, Klenow fragment of E.coli DNA polymerase, and reverse transcriptase.

A preferred amplification method is PCR. However, PCR conditions used for each reaction are empirically determined. A number of parameters influence the success of a reaction. Among them are annealing temperature and time, extension time, Mg^{2+} ATP concentration, pH, and the relative concentration of primers, templates, and deoxyribonucleotides. After amplification, the resulting DNA fragments can be detected by agarose gel electrophoresis followed by visualization with ethidium bromide staining and ultraviolet illumination.

The invention further provides the isolated polynucleotide operatively linked to a promoter of RNA transcription, as well as other regulatory

sequences for replication and/or transient or stable expression of the DNA or RNA. As used herein, the term "operatively linked" means positioned in such a manner that the promoter will direct transcription of RNA off the DNA molecule. Examples of such promoters are SP6, T4 and T7. In certain embodiments, cell-specific promoters are used for cell-specific expression of the inserted polynucleotide. Vectors which contain a promoter or a promoter/enhancer, with termination codons and selectable marker sequences, as well as a cloning site into which an inserted piece of DNA can be operatively linked to that promoter are well known in the art and commercially available. For general methodology and cloning strategies, see Gene Expression Technology (Goeddel ed., Academic Press, Inc. (1991)) and references cited therein and Vectors: Essential Data Series (Gacesa and Ramji, eds., John Wiley & Sons, N.Y. (1994)), which contains maps, functional properties, commercial suppliers and a reference to GenEMBL accession numbers for various suitable vectors. Preferable, these vectors are capable of transcribing RNA in vitro or in vivo.

Fragment of the sequences shown in SEQ ID NOS:1-732 or their respective complements also are encompassed by this invention, preferably at least 10 nucleotides and more preferably having at least 18 nucleotides. Larger polynucleotides, e.g., cDNA or genomic DNA, which hybridize under moderate or stringent conditions to the polynucleotide sequences shown in SEQ ID NOS:1-732, or their respective complements, also are encompassed by this invention.

In one embodiment, these fragments are polynucleotides that encode polypeptides or proteins having diagnostic and therapeutic utilities as described herein as well as probes to identify transcripts of the protein which may or may not be present. These nucleic acid fragments can by prepared, for example, by restriction enzyme digestion of the polynucleotide of SEQ ID NOS:1-732, or their complements, and then labeled with a detectable marker. Alternatively, random fragments can be generated using nick translation of the molecule. For

methodology for the preparation and labeling of such fragments, see Sambrook et al., (1989) supra.

Expression vectors containing these nucleic acids are useful to obtain host vector systems to produce proteins and polypeptides. It is implied that these expression vectors must be replicable in the host organisms either as episomes or as an integral part of the chromosomal DNA. Suitable expression vectors include viral vectors, including adenoviruses, adeno-associated viruses, retroviruses, cosmids, etc. Adenoviral vectors are particularly useful for introducing genes into tissues in vivo because of their high levels of expression and efficient transformation of cells both in vitro and in vivo. When a nucleic acid is inserted into a suitable host cell, e.g., a procaryotic or a eucaryotic cell and the host cell replicates, the protein can be recombinantly produced. Suitable host cells will depend on the vector and can include mammalian cells, animal cells, human cells, simian cells, insect cells, yeast cells, and bacterial cells constructed using well known methods. See Sambrook et al. (1989) supra. In addition to the use of viral vector for insertion of exogenous nucleic acid into cells, the nucleic acid can be inserted into the host cell by methods well known in the art such as transformation for bacterial cells; transfection using calcium phosphate precipitation for mammalian cells; or DEAE-dextran; electroporation; or microinjection. See Sambrook et al. (1989) supra for this methodology. Thus, this invention also provides a host cell, e.g. a mammalian cell, an animal cell (rat or mouse), a human cell, or a procaryotic cell such as a bacterial cell, containing a polynucleotide encoding a protein or polypeptide or antibody.

When the vectors are used for gene therapy *in vivo* or *ex vivo*, a pharmaceutically acceptable vector is preferred, such as a replication-incompetent retroviral or adenoviral vector. Pharmaceutically acceptable vectors containing the nucleic acids of this invention can be further modified for transient or stable expression of the inserted polynucleotide. As used herein, the term "pharmaceutically acceptable vector" includes, but is not limited to, a vector or delivery vehicle having the ability to selectively target

and introduce the nucleic acid into dividing cells. An example of such a vector is a "replication-incompetent" vector defined by its inability to produce viral proteins, precluding spread of the vector in the infected host cell. An example of a replication-incompetent retroviral vector is LNL6 (Miller, A.D. et al. (1989) BioTechniques 7:980-990). The methodology of using replication-incompetent retroviruses for retroviral-mediated gene transfer of gene markers is well established (Correll et al. (1989) PNAS USA 86:8912; Bordignon (1989) PNAS USA 86:8912-52; Culver, K. (1991) PNAS USA 88:3155; and Rill, D.R. (1991) Blood 79(10):2694-700. Clinical investigations have shown that there are few or no adverse effects associated with the viral vectors, see Anderson (1992) Science 256:808-13.

Compositions containing the polynucleotides of this invention, in isolated form or contained within a vector or host cell are further provided herein. When these compositions are to be used pharmaceutically, they are combined with a pharmaceutically acceptable carrier.

This invention further encompasses genes, either genomic or cDNA, which code for a polypeptide or protein in the cell of interest. The genes specifically hybridize under moderate or stringent conditions to a polynucleotide identified by SEQ ID NOS: 1-732 or their respective complements. The process of identification of larger fragment or the full-length coding sequence to which the partial sequence depicted in SEQ ID NOS:1-732 hybridizes preferably involves the use of the methods and reagents provided in this invention, either singularly or in combination.

Five methods are disclosed herein which allows one of skill in the art to isolate the gene or cDNA corresponding to the transcripts of the invention.

RACE-PCR Technique

One method to isolate the gene or cDNA which code for a polypeptide or protein and which corresponds to a transcript of this invention, involves the 5'-RACE-PCR technique. In this technique, the poly-A mRNA that contains the coding sequence of particular interest is first identified by hybridization to

a sequence disclosed herein and then reverse transcribed with a 3'-primer comprising the sequence disclosed herein. The newly synthesized cDNA strand is then tagged with an anchor primer of a known sequence, which preferably contains a convenient cloning restriction site attached at the 5'end. The tagged cDNA is then amplified with the 3'-primer (or a nested primer sharing sequence homology to the internal sequences of the coding region) and the 5'-anchor primer. The amplification may be conducted under conditions of various levels of stringency to optimize the amplification specificity. 5'-RACE-PCR can be readily performed using commercial kits (available from, e.g., BRL Life Technologies Inc, Clotech) according to the manufacturer's instructions.

Identification of known genes or ESTs

In addition, databases exist that reduce the complexity of ESTs by assembling contiguous EST sequences into tentative genes. For example, TIGR has assembled human ESTs into a datable called THC for tentative human consensus sequences. The THC database allows for a more definitive assignment compared to ESTs alone. Software programs exist (give examples) that allow for assembling ESTs into contiguous sequences from any organism.

Isolation of cDNAs from a library by probing with the SAGE transcript or tag

Alternatively, mRNA from a sample preparation was used to construct cDNA library in the ZAP Express vector following the procedure described in Velculescu et al. (1997) Science 270:484. The ZAP Express cDNA synthesis kit (Stratagene) was used accordingly to the manufacturer's protocol. Plates containing 250 to 2000 plaques are hybridized as described in Rupert et al. (1988) Mol. Cell. Bio. 8:3104 to oligonucleotide probes with the same conditions previously described for standard probes exxcept that the hybridization temperature is reduced to room temperature. Washes are performed in 6X standard-saline-citrate 0.1% SDS for 30 minutes at room temperature. The probes are labeled with 32P-ATP through use of T4 polynucletoide kinase.

Table 2 - Transcripts increased in colon cancer

Transcripts increased in only colon primary tumors compared to normal colon (61 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag Sequence	Tag_Number	NC	TU	CL	PT	PC	Accession	Gene Name
1	CATGCACCTAATTGG	H285759	612	755	411	161	333	F15516	H.sapiens mitochondrial EST sequence (1-t-12) from
2	CATGTGATTTCACTT	H933704	452	595	235	80	314	U35430	Human cytochrome c oxidase subunit III (COIII) pse
3	CATGCCTGTAATCCC	H388150	433	549	380	443	197	Z70701	H.sapiens mRNA (fetal brain cDNA c2_11).
								X71347	H.sapiens HNF1-C mRNA.
								X71346	H.sapiens HNF1-B mRNA.
4	CATGCACTACTCACC	H291282	293	527	78	14	83	U09500	Human mitochondrion cytochrome b gene, partial cds
5	CATGGTGAAACCCCA(G)	H753750	392	517	389	453	194	X66785	H.sapiens mRNA for transacylase (DBT).
								X17648	Human mRNA for granulocyte-macrophage colony-stimu
								U09087	Human thymopoietin beta mRNA, complete cds.
								U09088	Human thymopoietin gamma mRNA, complete cds.
								U20770	Human metastasis suppressor (KAI1) mRNA, complete
6	CATGGGCTTTAGGGA	H687915	37	372	6	29	11	W15552	zb91h11.s1 Soares parathyroid tumor NbHPA Homo sap
								W32091	zc05d03.s1 Soares parathyroid tumor NbHPA Homo sap
								R62866	yi11d07.r1 Homo sapiens cDNA clone 138925 5'.
7	CATGACTTTCCAAA	H130369	32	272	32	23	20	X89839	H.sapiens mitochondrial DNA for loop attachment se
8	CATGTGGTGTATGCA	H965434	53	271	6	30	5	T11555	A1486F Homo sapiens cDNA clone A1486 similar to Mi
9	CATGAGGGTGTTTTC	H175872	26	218	7	20	10	T15773	IB1870 Homo sapiens cDNA 3'end similar to Human mi
10	CATGAGGTCAGGAGA(T)	H177315	93	213	113	148	58	X12544	Human mRNA for HLA class II DR-beta (HLA-DR B).
								S73483	phosphorylase kinase catalytic subunit PHKG2 homol
11	CATGTTGGCCAGGCT	H1025322	124	194	63	111	51	X74301	H.sapiens mRNA for MHC class II transactivator.
								U28687	Human zinc finger containing protein ZNF157 (ZNF15
								U29119	Human leiomyoma LM-196.4 ectopic sequence from HMG
								U56236	Human Fc alpha receptor b mRNA, complete cds.
12	CATGATCACGCCCTC	H214616	97	186	17	41	49	W03751	za62h11.r1 Soares fetal liver spleen 1NFLS Homo sa
								W03770	za63f10.r1 Soares fetal liver spleen 1NFLS Homo sa

								W04748	za42f09.r1 Soares fetal liver spleen 1NFLS Homo sa
13	CATGGGGGTCAGGGG	H699691	37	170	11	16	9	T12078	A730R Homo sapiens cDNA clone A730 similar to Mito
								W45641	zc26a12.s1 Soares senescent fibroblasts NbHSF Homo
14	CATGGCTAGGTTTAT	H641789	38	144	13	25	13	D51017	Human fetal brain cDNA 3'-end GEN-007C04.
								D53694	Human fetal brain cDNA 3'-end GEN-117E01.
15	CATGCCCCGTACATC	H350996	56	132	35	0	18		Unknown
16	CATGAGTAGGTGGCC	H183018	18	131	2	17	7	D51021	Human fetal brain cDNA 3'-end GEN-007D07.
								D51052	Human fetal brain cDNA 3'-end GEN-009C05.
								D52836	Human fetal brain cDNA 3'-end GEN-089E01.
17	CATGCCTGTAGTCCC	H388278	79	124	61	71	23	D83195	Human DNA for Deoxyribonuclease I precursor.
18	CATGAGACCCACAAC	H136465	64	121	28	24	15	D54113	Human fetal brain cDNA 5'-end GEN-129B05.
19	CATGCATTTGTAATA	H327364	49	107	35	7	40	F15796	H.sapiens mitochondrial EST sequence (102-25) from
20	CATGTCCCCGTACCT	H874182	28	78	14	0	13		
21	CATGGCCAACCTCCT	H606582	23	73	8	6	19	Z59183	H.sapiens CpG island DNA genomic Mse1 fragment, cl
								D52905	Human fetal brain cDNA 5'-end GEN-091D11.
22	CATGGCCATCCCCTT	H609624	29	73	7	14	16	F16449	H.sapiens mitochondrial EST sequence (129-09) from
23	CATGTTGGTCAGGCT	H1027370	35	67	18	35	14	U06452	Human melanoma antigen recognized by T-cells (MART
24	CATGTCCTATTAAG	H881603	20	49	17	15	26		
25	CATGTTACTTATACT	H991026	2	47	2	1	4	D51004	Human fetal brain cDNA 3'-end GEN-006D02.
								L49057	Homo sapiens retinal fovea EST HFD010904 sequence.
								D51071	Human fetal brain cDNA 3'-end GEN-010E01.
26	CATGATGGCAGGAGT	H238755	13	45	1	4	2		
27	CATGCTAAGGCGAGG	H461411	5	44	2	3	3		
28	CATGGGTGAGACACT	H713234	7	44	20	13	15	J03592	Human ADP/ATP translocase mRNA, 3' end, clone pHAT
29	CATGACCTGTATCCC	H97078	6	42	17	100	32	X57352	Human 1-8U gene from interferon-inducible gene fam
30	CATGCCAGTCCGCCT	H339302	0	39	0	1	0	H01571	yj33e06.r1 Homo sapiens cDNA clone 150562 5' simil
								H03072	yj46g12.r1 Homo sapiens cDNA clone 151846 5' simil
31	CATGTAATTTTTGCC	H802810	1	37	0	1	0	T25155	EST730 Homo sapiens cDNA clone 34C11.
32	CATGTTAGCTTGTTT	H993264	6	37	2	3	5	D50972	Human fetal brain cDNA 3'-end GEN-004A05.
								D51211	Human fetal brain cDNA 3'-end GEN-017E08.
								D52162	Human fetal brain cDNA 3'-end GEN-069F04.
								T23865	seq2012 Homo sapiens cDNA clone Cot1374Ft-4HB3MA-3
33	CATGGCCACCCCCTG	H607576	0	35	1	0	0	M32053	Human H19 RNA gene, complete cds.
34	CATGTAATAAAGGTG	H798764	13	35	19	33	51	X67247	H.sapiens rpS8 gene for ribosomal protein S8.
35	CATGTACTGCTCGGA	H817627	13	35	5	1	14	T11939	A953F Homo sapiens cDNA clone A953 similar to Mito

36	CATGGTGAAACCCA	H753749	9	31	22	30	4	T95857	ye42f01.s1 Homo sapiens cDNA clone 120409 3' simil
								W03237	za35b09.r1 Soares fetal liver spleen 1NFLS Homo sa
								W03326	za63g03.r1 Soares fetal liver spleen 1NFLS Homo sa
37	CATGGAAACTGAACA	H526210	6	26	17	5	3	X54195	Human line-1 element DNA, host sequence flanking t
								U29607	Human methionine aminopeptidase mRNA, complete cds
								H95100	yw57b10.r1 Homo sapiens cDNA clone 256315 5' simil
38	CATGACTTTTTAAAA	H131009	1	22	4	1	0		
39	CATGGACTGCGTGCC	H555450	0	21	7	9	12	D29062	Human keratinocyte cDNA, clone 067.
								D29563	Human keratinocyte cDNA, clone 713.
								T03196	FB3B5 Homo sapiens cDNA clone FB3B5 3'end.
40	CATGTCAGTGGTAGT	H863923	4	21	2	2	1	Z57093	H.sapiens CpG DNA, clone 164a10, reverse read cpg1
41	CATGAAACTGTGGTT	H7916	2	20	2	2	1	Z60184	H.sapiens CpG island DNA genomic Mse1 fragment, cl
								Z63649	H.sapiens CpG island DNA genomic Mse1 fragment, cl
								W31349	zb95d06.s1 Soares parathyroid tumor NbHPA Homo sap
42	CATGGGGGGGGGGGT	H699051	0	19	0	0	0		
43	CATGGTGCCCGTGCC		2	19	1	0	0	W31448	zb96h01.s1 Soares parathyroid tumor NbHPA Homo sap
								W47282	zc40b06.r1 Soares senescent fibroblasts NbHSF Homo
44	CATGGGGGGTAACTA	H699144	3	19	15	12	5	X71428	H.sapiens fus mRNA.
								S62140	TLS=translocated in liposarcoma [human, mRNA, 1824
								W31782	zb96a06.r1 Soares parathyroid tumor NbHPA Homo sap
45	CATGTCCTGCCCCAT	H883029	3	19	14	27	16	M24398	Human parathymosin mRNA, complete cds.
46	CATGAAGTGGCAAGA	H47683	0	16	0	0	0		
47	CATGGGTATTAACCA	H708358	0	16	0	0	0	U33317	Human defensin 6 (HD-6) gene, complete cds.
								M98331	Homo sapiens defensin 6 mRNA, complete cds.
48	CATGGGCTACACCTT	H684312	2	16	0	2	1	D32027	Human mRNA for T cell receptor V beta 14 CDR3, par
			2	16	0	2	1	T11701	A1225F Homo sapiens cDNA clone A1225 similar to Mi
49	CATGAGGGTGTTTCC	H175870	1	15	0	0	0	D51783	Human fetal brain cDNA 5'-end GEN-051G02.
50	CATGCAAGGACCAGC	H272467	0	13	0	2	0	D13138	Human mRNA for dipeptidase.
									Homo sapiens (clones MDP4, MDP7) microsomal dipept
									RDP=renal dipeptidase [human, kidney, Genomic, 357
51	CATGTGGAAATGACC	H950498	0	13	0	167	0	M10629	Human alpha-1 collagen gene, 3' end with polyA sit
52	CATGATCCGCCTGCC	H219514	1	13	3	4	1	H11641	ym17e04.s1 Homo sapiens cDNA clone 47962 3' simila
								R95667	yq51a09.s1 Homo sapiens cDNA clone 199288 3' simil
53	CATGTCCCGTACAC	H875282	1	13	0	0	1		
54	CATGATGTAAAAAAT	H241665	0	11	0	12	14	M74090	Human TB2 gene mRNA, 3' end.

								J03801	Human lysozyme mRNA, complete cds with an Alu repe
								M19045	Human lysozyme mRNA, complete cds.
55	CATGCCAGCCCCGTC	H337244	0	11	0	0	0		
56	CATGACCATTCTGCT	H85882	0	10	1	26	3	X57351	Human 1-8D gene from interferon-inducible gene fam
								X02490	Human interferon-inducible mRNA (cDNA 1-8).
57	CATGAGGACCATCGC	H165175	0	10	0	0	0		
58	CATGATGTGAAGAGT(A)	H243747	0	10	0	165	0	J03040	Human SPARC/osteonectin mRNA, complete cds.
59	CATGCAGTTGGTTGT	H310975	0	10	6	7	4	U55217	Human RNA fragment from patients with Crohn's dise
60	CATGGCCCTCTGCCA	H613862	0	10	2	15	7		
61	CATGTTAGATAAGCA	H992010	0	10	3	3	6	M94083	Human chaperonin-like protein (HTR3) mRNA, complet
								L27706	Human chaperonin protein (Tcp20) gene complete cds

Transcripts increased in both colon primary tumors and colon cancer cell lines compared to normal colon (47 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag Sequence	Tag_Number	NC	CT	CL	PT	PC	Accession	Gene Name
1	CATGGCAGCCATCCG	H599350	87	180	230	72	138	U14969	Human ribosomal protein L28 mRNA, complete cds.
2	CATGATGGCTGGTAT	H239533	52	153	318	80	294	X17206	Human mRNA for LLRep3.
3	CATGCCCGTCCGGAA	H355689	87	142	246	178	250	X64707	H.sapiens BBC1 mRNA
4	CATGAGGCTACGGAA	H171113	44	117	167	86	147	X56932	H.sapiens mRNA for 23 kD highly basic protein
5	CATGAGCACCTCCAG	H148949	42	116	197	103	190	Z11692	H.sapiens mRNA for elongation factor 2.
6	CATGCTGGGTTAATA	H502724	29	115	160	75	134	M81757	H.sapiens S19 ribosomal protein mRNA, complete cds
7	CATGGGATTTGGCCT	H671654	55	108	222	73	185	M17887	Human acidic ribosomal phosphoprotein P2 mRNA, com
8	CATGTACCATCAATA	H807748	46	107	98	64	189	X53778	H.sapiens hng mRNA for uracil DNA glycosylase.
								J02642	Human glyceraldehyde 3-phosphate dehydrogenase mRN
9	CATGTGGGCAAAGCC	H959498	51	103	156	45	152	Z11531	H.sapiens mRNA for elongation factor-1-gamma.
								M55409	Human pancreatic tumor-related protein mRNA, 3' en
10	CATGAATCCTGTGGA	H55227	30	95	102	48	156	Z28407	H.sapiens mRNA for ribosomal protein L8.
11	CATGGGACCACTGAA	H660601	36	92	114	43	63	X73460	H.sapiens mRNA for ribosomal protein L3.
12	CATGAGGGCTTCCAA	H174037	47	91	167	91	155	M73791	Human novel gene mRNA, complete cds.
								M64241	Human Wilm's tumor-related protein (QM) mRNA, comp
								S35960	laminin receptor homolog {3' region} [human, mRNA
13	CATGAAGGTGGAGGA	H44683	48	91	182	113	215	X80822	H.sapiens mRNA for ORF.
14	CATGTGCACGTTTTC	H935680	45	87	105	61	122	X03342	Human mRNA for ribosomal protein L32
15	CATGTCAGATCTTTG	H861056	37	81	93	50	92	M58458	Human ribosomal protein S4 (RPS4X) isoform mRNA, c
								M22146	Human scar protein mRNA, complete cds.
16	CATGTGGTGTTGAGG	H965603	42	79	83	55	250	X69150	H.sapiens mRNA for ribosomal protein S18.
								L06432	Homo sapiens 18S ribosomal protein (HKE3) mRNA seq
17	CATGCCTAGCTGGAT	H379369	28	77	80	46	143	Y00052	Human mRNA for T-cell cyclophilin.
18	CATGCTTGGGTTTTG	518912	0	73	42	0	0	X07868	Human DNA for insulin-like growth factor II (IGF-2);
19	CATGCTCCTCACCTG	H482584	12	72	41	34	50	U16811	Human Bak mRNA, complete cds.

20	CATGCTGTTGGTGAT	H507577	17	65	116	48	103	D14530	Human homolog of yeast ribosomal protein S28, comp
21	CATGCGCCGGAACAC	H416261	28	62	183	55	94	X73974	H.sapiens HRPL4 mRNA.
22	CATGCAATAAATGTT	H274492	9	60	73	55	119	D23661	Human mRNA for ribosomal protein L37, complete cds
23	CATGACATCATCGAT	H79065	15	57	82	42	118	L06505	Human ribosomal protein L12 mRNA, complete cds.
24	CATGTTCAATAAAAA	H1000193	12	56	154	49	99	M17886	Human acidic ribosomal phosphoprotein P1 mRNA, com
25	CATGGAACACATCCA	H528694	24	56	71	24	146	X63527	H.sapiens mRNA for ribosomal protein L19.
26	CATGTTATGGGATCT	H998030	7	55	78	35	77	M24194	Human MHC protein homologous to chicken B complex
27	CATGGCATAATAGGT		18	53	50	19	61	U14967	Human ribosomal protein L21 mRNA, complete cds.
28	CATGATTCTCCAGTA	H253260	23	50	103	49	120	X55954	Human mRNA for HL23 ribosomal protein homologue.
								X52839	Human mRNA for ribosomal protein L17.
29	CATGACTCCAAAAAA	H119809	15	49	64	21	64	H38868	yp61a04.r1 Homo sapiens cDNA clone 191886 5' simil
								H71935	ys15f12.r1 Homo sapiens cDNA clone 214895 5'.
								Z43914	H. sapiens partial cDNA sequence; clone c-1od03.
								T48545	hbc3221 Homo sapiens cDNA clone hbc3221 5'end.
30	CATGCTGTTGATTGC	H507455	9	44	54	22	40	X04347	Human liver mRNA fragment DNA binding protein UP1
31	CATGTACAAAATCGA	802871	0	42	20	0	0	X00910	Human mRNA for IGF-II precursor (insulin-like grow
32	CATGGAAAAATGGTT	H524524	14	41	81	15	57	X61156	H.sapiens mRNA for laminin-binding protein.
								J03799	Human colin carcinoma laminin-binding protein mRNA
33	CATGAAGAAGATAGA	H33331	9	39	69	30	56	U02032	Human ribosomal protein L23a mRNA, partial cds.
34	CATGCCTTCGAGATC	H390692	12	36	51	25	86	U14970	Human ribosomal protein S5 mRNA, complete cds.
35	CATGACTGGGTCTAT	H125661	5	29	25	25	38	X58965	H.sapiens RNA for nm23-H2 gene.
								M36981	Human putative NDP kinase (nm23-H2S) mRNA, complet
								L16785	Homo sapiens c-myc transcription factor (puf) mRNA
36	CATGCAGCTCACTGA	H302367	9	29	40	27	31	L10376	Human (clone CTG-B33) mRNA sequence.
								S80520	CAG-isl 7 {trinucleotide repeat-containing sequenc
37	CATGGTGTGTTTGTA	H769020	0	24	15	22	8	M77349	Human transforming growth factor-beta induced gene
38	CATGGTGCGCTGAGC	H760291	0	22	17	44	18	X58536	Human mRNA for HLA class I locus C heavy chain.
39	CATGGTTCACATTAG	H774461	3	22	25	141	10	X00497	Human mRNA for HLA-DR antigens associated invarian
40	CATGTGAAATAAAAC	H918273	2	18	37	8	22	X16934	Human hB23 gene for B23 nucleophosmin.
41	CATGAAAAGAAACTT	H2056	1	16	27	11	25	Y00345	Human mRNA for polyA binding protein.
42	CATGTGCTGCCTGTT	H948604	1	15	16	11	3	X81005	H.sapiens HCG IV mRNA.
								D28137	Human mRNA for BST-2, complete cds.
43	CATGCTGATGGCAGA	H495251	0	14	15	8	6	W46476	Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 324128 3'.
								X72718	H.sapiens DNA for orphan TCR V-beta segment (allel

44	CATGACTCGCTCTGT	H121311	0	12	16	5	7	H121311	Soares fetal heart NbHH19W Homo sapiens cDNA clone 342926 3'.
								AA305589	EST176663 Colon carcinoma (Caco-2) cell line II Homo sapiens cDNA 5' end
45	CATGGCCCAAGGACC	H610466	0	12	19	82	17	X53416	Human mRNA for actin-binding protein (filamin) (AB
46	CATGATCTTGTTACT	H229106	0	11	28	67	0	X02761	Human mRNA for fibronectin (FN precursor).
47	CATGAAGCTGCTGGA	H40571	0	10	17	6	6	Z26305	H.sapiens isoform 1 gene for L-type calcium channe

Transcripts increased in only colon cancer cell lines compared to normal colon (181 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag Sequence	Tag_Number	NC	TU	CL	PT	PC	Accession	Gene Name
1	CATGTGTGTTGAGAG	H978825	71	79	487	136	412	X16869	Human mRNA for elongation factor 1-alpha
2	CATGGCCGAGGAAGG	H615043	72	66	265	105	125	X53505	Human ribosomal protein S12.
3	CATGCAAACCATCCA	H263478	137	83	245	36	502	X12883	Human cytokeratin 18.
4	CATGCACAAACGGTA	H278636	63	53	201	74	179	L19739	Homo sapiens metallopanstimulin (MPS1)
5	CATGAAAAAAAAAAA	H1	31	48	186	66	102	X83412	H.sapiens B1 mRNA for mucin.
								Z32564	H.sapiens FRGAMMA mRNA (819bp) for folate receptor
								X76180	H.sapiens mRNA for lung amiloride sensitive Na+ ch
								U08470	Human FR-gamma' mRNA, complete cds.
								U08471	Human folate receptor 3 mRNA, complete cds.
6	CATGTTGGTCCTCTG	H1027448	115	128	179	104	358	S64030	Human L41 ribosomal protein
7	CATGTCTCCATACCC	H906438	0	0	176	48	0	T91925	ye02f02.r1 Homo sapiens cDNA clone 116571 5'.
8	CATGAAGACAGTGGC	H33979	59	61	172	55	252	X66699	H.sapiens ribosomal protein L37a.
9	CATGCCGTCCAAGGG	H374027	50	39	138	60	108	M60854	Human ribosomal protein S16
10	CATGGGGGAAATCGC	H696375	90	90	136	203	231	M92381	Human thymosin beta 10
11	CATGAAGGAGATGGG	H41531	30	37	133	38	161	X69181	H.sapiens mRNA for ribosomal protein L31.
12	CATGGAGGGAGTTTC	H567488	38	53	112	65	142	U14968	Human ribosomal protein L27a
13	CATGCGCTGGTTCCA	H424694	42	64	111	53	49	X79234	H.sapiens ribosomal protein L11.
14	CATGGCCGTGTCCGC	H618199	56	39	109	28	120	J03537	Human ribosomal protein S6
15	CATGGACGACACGAG	H549145	32	59	105	44	70	U58682	Human ribosomal protein S28 mRNA, complete cds
16	CATGTCACCCACACC	H857362	36	48	103	44	65	X52839	Human mRNA for ribosomal protein L17
17	CATGCGCCGCCGGCT	H416106	39	43	90	52	184	U12465	Human ribosomal protein L35
18	CATGCTCAACATCTC	H475448	27	41	89	27	145	M17885	Human acidic ribosomal phosphoprotein P0
19	CATGTGGCCCCACCC	H955718	20	30	80	46	55	M23725	Human M2-type pyruvate kinase mRNA, complete cds.
								M26252	Human TCB gene encoding cytosolic thyroid hormone-
20	CATGCCCTGGGTTCT	H359102	34	49	78	92	145	M11147	Human ferritin L chain

21	CATGAGCATCTCCAG	H150997	0	0	77	0	0	H09058	yl96f11.r1 Homo sapiens cDNA clone 45943 5'.
								Z44640	H. sapiens partial cDNA sequence; clone c-26b05.
								N75111	yz29e01.r1 Homo sapiens cDNA clone 284472 5'.
22	CATGGCCTGTATGAG	H621369	24	32	77	33	99	M31520	Human ribosomal protein S24 mRNA.
23	CATGAGCTCTCCCTG	H161624	33	39	76	21	67	X53777	Human L23 mRNA for putative ribosomal protein.
24	CATGCCAGGAGGAAT	H338081	27	12	74	23	87	AA223340	gb|AA223340|AA223340 Homo sapiens cDNA clone 650651 3' similar to gb:Y00371_rna1 HEAT SHOCK COGNATE 71 KD PROTEIN (HUMAN)
25	CATGGGCAAGCCCCA	H672342	30	55	72	27	61	U12404	Human Csa-19
26	CATGAGGAAAGCTGC	H163999	31	42	70	32	146	F16378	H.sapiens EST sequence (135-18) from skeletal muscle
27	CATGAACGCGGCCAA	H26261	29	46	69	54	79	Z23063	Homo sapiens macrophage migration inhibitory factor
28	CATGCCAGAACAGAC	H335945	23	39	66	42	148	X79238	H.sapiens ribosomal protein L30.
29	CATGGCCGCCATCTC	H615736	7	10	65	10	22	U55017	Human transketolase (TKT)
30	CATGGTGTTAACCAG	H769045	16	19	65	17	76	L25899	Human ribosomal protein L10
31	CATGCCTCGGAAAAT	H383489	9	13	64	23	46	Z26876	H.sapiens ribosomal protein L38.
32	CATGAGGTCCTAGCC	H177610	15	27	63	43	41	X06547	Human class Pi glutathione S-transferase
33	CATGGTTCCCTGGCC	H775658	31	26	63	32	96	X65923	H.sapiens fau mRNA.
34	CATGTAAGGAGCTGA	H796831	32	58	62	42	68	X77770	H.sapiens RPS26
35	CATGAACTAAAAAAA	H28673	7	14	60	17	39	W52460	zc45e11.r1 Soares senescent fibroblasts NbHSF Homo
								N92893	zb71h03.s1 Homo sapiens cDNA clone 309077 3'.
36	CATGATTTGTCCCAG	H260949	17	13	57	9	91	X14957	Human hmgI mRNA for high mobility group protein I.
37	CATGATAATTCTTTG	H200576	13	27	53	30	69	U14973	Human ribosomal protein S29
38	CATGCCCCAGCCAGT	H348756	18	23	53	5	85	U14990	Human XP1PO ribosomal protein S3 (rpS3)
39	CATGGGAGTGGACAT	H667269	15	13	49	13	45	L11566	Homo sapiens ribosomal protein L18 (RPL18)
40	CATGTAAAAAAAAAA	H786433	13	8	48	10	26	H08238	yl87a01.r1 Homo sapiens cDNA clone 44932 5'.
41	CATGGTGTTGCACAA	H769605	19	21	48	21	47	X79239	H.sapiens ribosomal protein S13.
42	CATGGCCAGCCCAGC	H608595	6	21	47	11	15	U31657	Human unknown protein mRNA, partial cds.
								H41030	yn92a10.r1 Homo sapiens cDNA clone 175866 5'.
43	CATGGGCTCCCACTG	H685384	14	24	47	23	15	M16660	Human 90-kDa heat-shock protein
44	CATGTCAACTTCTGG	H853983	0	0	46	2	0	N57419	yw82e04.r1 Homo sapiens cDNA clone 258750 5' simil
45	CATGGATGCTGCCAA	H583573	6	12	46	27	18	X59357	Human mRNA for Epstein-Barr virus small RNAs (EBER)
								L21756	Homo sapiens acute myeloid leukemia associated protein
								D17652	Human mRNA for HBp15/L22, complete cds.
46	CATGAATAGGTCCAA	H51925	13	31	46	47	53	M64716	Human ribosomal protein S25
47	CATGGCTTTTAAGGA	H655115	8	26	45	22	63	L06498	Homo sapiens ribosomal protein S20 (RPS20)
48	CATGAATGCAGGCAG	H58533	2	12	44	6	27	M61831	Human S-adenosylhomocysteine hydrolase (AHCY)

49	CATGGCCCAGCTGGA	H610939	8	18	43	0	22	Z21507	Human elongation factor 1 delta (EF 1delta)
50	CATGGGCCGCGTTCG	H678334	6	6	42	8	18	M13932	Human ribosomal protein S17 mRNA
51	CATGTGAGGGAATAA	H928269	14	26	42	15	42	M10036	Human triosephosphate isomerase
52	CATGTGTACCTGTAA	H968173	14	24	42	35	49	K00558	human alpha-tubulin
53	CATGGGCAAGAAGAA	H672265	8	7	41	12	87	L19527	Homo sapiens ribosomal protein L27 (RPL27)
54	CATGAACTAACAAAA	H28737	6	14	40	14	15	X63237	H.sapiens Uba80 mRNA for ubiquitin.
55	CATGTATACGCTCAG	H837237	0	0	38	0	9		Unknown
56	CATGTACAAGAGGAA	H803369	7	17	38	14	42	X69391	H.sapiens ribosomal protein L6.
57	CATGGTTAACGTCCC	H770486	8	17	38	12	25	H11182	ym14a02.r1 Homo sapiens cDNA clone 47866 5'
								T40302	ya31g04.r5 Homo sapiens cDNA clone 62262 5'
								T89480	yd98a05.r1 Homo sapiens cDNA clone 116240 5'
58	CATGGAGACTCCTGC	H558943	13	12	38	32	10	H01362	yi99c06.r1 Homo sapiens cDNA clone 147370 5'
59	CATGATCCACATCGC	H217399	3	10	37	10	14	H94371	yw54e05.r1 Homo sapiens cDNA clone 256064 5'.
								T49412	ya75b09.r1 Homo sapiens cDNA clone 67481 5'.
								T51058	yb55a12.r1 Homo sapiens cDNA clone 75070 5'.
60	CATGGAAGCTTTGCA	H534522	11	13	37	14	25	X07270	Human heat shock protein hsp86.
61	CATGCTGGCGAGCGC	H501287	2	9	36	3	18	M91670	Human ubiquitin carrier protein (E2-EPF)
62	CATGCTGAGACAAAG	H493633	13	8	36	8	26	X74070	H.sapiens transcription factor BTF 3.
63	CATGAACGACCTCGT	H24951	7	13	35	22	40	V00599	Human beta-tubulin
64	CATGGCATAGGCTGC	H602783	9	16	35	2	17	X84694	H.sapiens mRNA for elongations factor Tu-mitochondria
								L38995	Homo sapiens nuclear-encoded mitochondrial elongatation factor
								S75463	P43=mitochondrial elongation factor homolog [human
65	CATGCATCTTCACCA	H319302	12	14	35	9	16	H48893	yq80b12.r1 Homo sapiens cDNA clone 202079 5'
66	CATGGCCTGCTGGGC	H621035	10	5	32	18	107	X71973	H.sapiens GPx-4 mRNA for phospholipid hydroperoxidase
67	CATGACAGGCTACGG	H76231	0	5	31	64	0	M95787	Human 22kDa smooth muscle protein (SM22)
68	CATGGAAATGTAAGA	H528067	5	12	31	14	25	H80294	yu59g01.s1 Homo sapiens cDNA clone 230448 3'.
								R74294	yi57f06.r1 Homo sapiens cDNA clone 143363 5'.
69	CATGGAAGCCAGCCA	H533798	1	3	30	9	11	L36055	Human 4E-binding protein 1
70	CATGTTACCATATCA	H988366	10	28	30	19	86	F17005	H.sapiens EST sequence (011-T1-18) from skeletal muscle
71	CATGTTGCTCACAAA	H1023249	1	2	29	1	2	H10519	yl90g04.r1 Homo sapiens cDNA clone 45563 5'.
72	CATGTCCCCGCTCGA	H874103	0	6	29	0	0		Unknown
73	CATGATTAACAAAGC	H246019	8	9	29	25	26	X04409	Human coupling protein G(s) alpha-subunit
74	CATGCAGATCTTTGT	H298495	2	7	28	8	24	X56998	Human UbA52 adrenal mRNA for ubiquitin-52 amino acid
75	CATGGTTCGTGCCAA	H777109	9	28	28	17	46	F19234	H.sapiens EST sequence (005-X3-16) from skeletal m
76	CATGGACGTGTGGGC	H552683	3	4	27	2	16	X52317	Human histone H2A.Z.

77	CATGCTAAAAAAAAA	H458753	4	8	27	19	8	M33680	Human 26-kDa cell surface protein TAPA-1
78	CATGGGGTTTTTATT	H704500	4	1	27	6	18	L28809	Homo sapiens dbpB-like protein
79	CATGCCGATCACCGG	H363799	7	9	27	7	15	M29536	Human translational initiation factor 2 beta subunit
80	CATGGCACAAGAAGA	H594051	6	9	26	7	29	W07137	za92a11.r1 Soares fetal lung NbHL19W Homo sapiens
								D20503	Human HL60 3'directed MboI cDNA, HUMGS01477, clone
								N91592	Soares fetal lung NbHL19W Homo sapiens cDNA clone 303055 3'.
								H83884	yv84c07.s1 Homo sapiens cDNA clone 249420 3' similar to contains Alu repetitive element;.
81	CATGTCTCTACCCAC	H908373	7	11	26	11	13	Z22572	H.sapiens CDEI binding protein mRNA.
								L09209	Homo sapiens amyloid protein homologue mRNA, compl
								L19597	Human binding protein mRNA, partial cds.
								S60099	APPH=amyloid precursor protein homolog [human, pla
82	CATGGTTTCCCCAAG	H783697	1	0	25	3	0	W07587	zb06f02.r1 Soares fetal lung NbHL19W Homo sapiens
								N28502	yx36f06.r1 Homo sapiens cDNA clone 263843 5'
								N35630	yx62a03.r1 Homo sapiens cDNA clone 266284 5'
83	CATGCCTGTCCAGCC	H388426	2	3	25	3	13	Z40265	H. sapiens partial cDNA sequence; clone c-1xe03.
								W02723	zc65c03.s1 Soares fetal heart NbHH19W Homo sapiens
								N24893	yx99h09.s1 Homo sapiens cDNA clone 269921 3'.
								N32178	yy25b09.s1 Homo sapiens cDNA clone 272249 3'.
84	CATGTCATCATCTGA	H865503	5	15	25	5	7	H21873	yl34b10.s1 Homo sapiens cDNA clone 160123 3' simil
								H26394	yl48e12.s1 Homo sapiens cDNA clone 161518 3' simil
								H69857	yr88d02.s1 Homo sapiens cDNA clone 212355 3' simil
								H70714	yu69b11.s1 Homo sapiens cDNA clone 239037 3' simil
85	CATGCCCTGCCTTGT	H358783	5	8	25	16	31	X55110	Human mRNA for neurite outgrowth-promoting protein
86	CATGGCCGGGCCCTC	H617048	1	1	24	0	1	X03168	Human mRNA for S-protein.
87	CATGTTGCTCAAAAA	H1023233	2	1	24	2	2	AA143561	zo32d09.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588593 3' similar to contains LTR7.t1 LTR7 repetitive element
								AA152342	zo01g11.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 566468 3' similar to contains LTR7.t3 LTR7 repetitive element ;
								AA115727	zl86h11.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511557 3' similar to contains LTR7.t1 LTR7 repetitive element
88	CATGCAAAATCAGGA	H262987	6	2	24	5	15	R76502	yi61f09.r1 Homo sapiens cDNA clone 143753 5'.
								T32681	EST52915 Homo sapiens cDNA 5' end similar to None.
								T34662	EST72468 Homo sapiens cDNA 5' end similar to None.
89	CATGGAAGATGTGGG	H533435	1	5	23	4	7	H04634	yj49h03.r1 Homo sapiens cDNA clone 152117 5'.

								F00364	H. sapiens partial cDNA sequence; clone 76D12; ver
90	CATGGTGCTCATTCA	H761150	0	8	23	6	4	H01503	yj21c05.s1 Homo sapiens cDNA clone 149384 3'.
								H84813	yv86c02.s1 Homo sapiens cDNA clone 249602 3' simil
								H84956	yv88f07.s1 Homo sapiens cDNA clone 249829 3' simil
91	CATGGCTTTACTTTG	H654464	4	5	23	9	5	L38961	Homo sapiens putative transmembrane protein (B5)
92	CATGTTTTCTGAAAA	H1046401	6	13	23	10	10	J04026	Human thioredoxin (TXN) mRNA
93	CATGTTGCTCACACA	H1023250	1	4	22	0	4	D11078	Human RGH2 gene.
94	CATGGATTTCTCAGC	H589267	0	0	22	0	19	X53279	Human mRNA for placental-like alkaline phosphatase
95	CATGAGGAGGGAGGC	H166539	2	3	22	2	4	M77836	Human pyrroline 5-carboxylate reductase mRNA,
96	CATGGCTTAACCTGG	H651359	3	4	22	2	4	X07674	Human glutamate dehydrogenase
97	CATGCTCTTCGAGAA	H490889	4	8	22	27	19	Y00433	Human mRNA for glutathione peroxidase
98	CATGAGAACAAAACC	H132098	1	7	21	9	6	X67951	H.sapiens mRNA for proliferation-associated gene
99	CATGCCCAGGGAGAA	H346761	3	3	21	2	24	U38846	Human stimulator of TAR RNA binding (SRB)
								D16933	Human HepG2 3' region cDNA, clone hmd4f11.
100	CATGCACTTCAAGGG	H294155	0	3	20	47	107	U42376	Human retinoic acid induced RIG-E
101	CATGGCGGAGAGAGG	H631331	2	3	20	4	1		Unknown
102	CATGTTACCTCCTTC	H989024	4	7	20	3	22	F17524	H.sapiens EST sequence (012-T2-32) from skeletal m
103	CATGACTCTGCCAAG	H122449	4	7	20	3	7		Unknown
104	CATGTCAGATGGCGT	H861095	1	6	19	12	7	W52942	zc03h05.r1 Soares parathyroid tumor NbHPA Homo sap
105	CATGGGCCTTTTTTT	H679936	1	3	19	5	3	R21316	yg48h11.r1 Homo sapiens cDNA clone 35917 5' simila
106	CATGTGGACGCGCTG	H951912	0	0	19	0	0	X00566	Human lipoprotein apoA1.
107	CATGCCTGCTCCCTG	H386904	0	5	19	6	5	M80244	Human E16 mRNA
108	CATGGCCACACCCCA(C)	H607318	2	6	18	18	15	H27927	yl58c11.s1 Homo sapiens cDNA clone 162452 3' simil
109	CATGATTATTTTTCT	H249854	2	3	18	5	20	X57959	H.sapiens ribosomal protein L7.
110	CATGGAACCCTGGGA	H529899	2	7	18	5	15	AA299898	EST12509 Uterus tumor I Homo sapiens cDNA 5' end
111	CATGGGCTGATGTGG	H686319	3	5	18	8	17	U09510	Human glycyl-tRNA synthetase.
112	CATGTCAATAAAGAA	H855049	3	10	18	4	4	X76013	H.sapiens QRSHs mRNA for glutaminyl-tRNA synthetas
113	CATGAAAGTGAAGAT	H11785	0	7	17	0	5	W16529	zb10a11.r1 Soares fetal lung NbHL19W Homo sapiens
								W35192	zc70b05.r1 Soares fetal heart NbHH19W Homo sapiens
								W52451	zc45d09.r1 Soares senescent fibroblasts NbHSF Homo
114	CATGCACGCGCTCAA	H288373	0	1	17	0	3	D38251	Human mRNA for RPB5 (XAP4)
115	CATGAACTAATACTA	H28872	1	6	17	13	31	D52570	Human fetal brain cDNA 5'-end GEN-081G12.
								D52758	Human fetal brain cDNA 5'-end GEN-087A08.
								D55953	Human fetal brain cDNA 5'-end GEN-407H12.
116	CATGCTGTACCTGGA	H504187	1	0	17	12	6	M22490	Human bone morphogenetic protein-2B (BMP-2B)

117	CATGCGACCCCACGC	H398663	2	6	17	48	0	M12529	Human apolipoprotein E
118	CATGTAGAAAAATAA	H819213	0	1	16	2	7	X16539	H.sapiens RNA for neuroleukin gene.
								M27691	Human transactivator protein (CREB) mRNA, complete
119	CATGATCTTGAAAGG	H228867	0	0	16	5	3	M86667	H.sapiens NAP (nucleosome assembly protein)
120	CATGCAGCTGGCCAT	H302741	0	1	16	14	0	X53743	H.sapiens mRNA for fibulin-1 C.
121	CATGATCTTGAAAGG	H228867	0	0	16	5	3	Z26328	H. sapiens partial cDNA sequence; clone HEC059
121	CATGATCTTGAAAGG	H228867	0	0	16	5	3	Z26328	H. sapiens partial cDNA sequence; clone HEC059
122	CATGGTGGAGGTGCG	H762554	2	10	16	3	5	U22055	Human 100 kDa coactivator mRNA
123	CATGGTGGACCCCAA	H762197	1	5	15	7	10	R91724	yp98e02.r1 Homo sapiens cDNA clone 195482 5' simil
								W51770	zc48a02.r1 Soares senescent fibroblasts NbHSF Homo
								N42086	yy05b03.r1 Homo sapiens cDNA clone 270317 5'
124	CATGGAGCAGCTGGA	H561787	0	5	15	2	4	R80990	yi94c02.r1 Homo sapiens cDNA clone 146882 5'
								R95056	yq44f01.r1 Homo sapiens cDNA clone 198649 5' simil
125	CATGGCGGGAGGGCT	H633002	1	6	15	8	7	F16507	H.sapiens EST sequence (147-09) from skeletal musc
								T50201	yb77h05.r1 Homo sapiens cDNA clone 77241 5' simila
126	CATGATTGGCTTAAA	H256497	1	8	15	0	16	S85655	Human prohibitin
127	CATGGAAAAATTTAA	H524541	0	3	15	4	0	M38188	Human unknown protein from clone pHGR74 mRNA, comp
128	CATGGATCACAGTTT	H577840	0	5	15	0	0	Y00711	Human lactate dehydrogenase B (LDH-B).
129	CATGAGCCTTTGTTG	H155632	1	2	15	23	5	D83174	Human collagen binding protein 2.
130	CATGTCTGCACCTCC	H910430	0	0	15	0	2	X70940	H.sapiens elongation factor 1 alpha-2.
131	CATGAACAGAAGCAA	H18469	0	2	15	3	11	T30623	EST19638 Homo sapiens cDNA 5' end similar to None.
								C01011	HUMGS0004747, Human Gene Signature, 3'-directed cDNA sequence.
								AA111865	zm62d06.s1 Stratagene fibroblast (#937212) Homo sapiens cDNA clone 530219 3'
								W56516	zd16c08.r1 Soares fetal heart NbHH19W Homo sapiens
132	CATGTGTTCAGGACC	H980130	1	1	14	5	11	H30299	yo77d04.r1 Homo sapiens cDNA clone 183943 5' simil
								H50265	yo28c02.r1 Homo sapiens cDNA clone 179234 5'.
133	CATGTAGATAATGGC	H822331	1	4	14	6	14	W01702	za37a06.r1 Soares fetal liver spleen 1NFLS Homo sa
								W04495	za58b10.r1 Soares fetal liver spleen 1NFLS Homo sa
								W23528	zc71g11.s1 Soares fetal heart NbHH19W Homo sapiens
134	CATGCTTAATCCTGA	H508767	0	6	14	6	12	D11838	Human HepG2 3'-directed MboI cDNA, clone hm02e09.
135	CATGGGCAGAGGACC	H673954	0	6	14	5	11	X75598	H.sapiens nm23H1 gene.
136	CATGTGACTGAAGCC	H925194	0	5	14	3	0	T35470	EST85850 Homo sapiens cDNA 5' end similar to None.
								T35536	EST86951 Homo sapiens cDNA 5' end similar to None.

								T35545	EST87066 Homo sapiens cDNA 5' end similar to None.
137	CATGGATAGTTGTGG	H576495	0	1	14	2	1	H01694	yj33g11.s1 Homo sapiens cDNA clone 150596 3'.
								N78851	zb17d08.s1 Homo sapiens cDNA clone 302319 3'.
								N78931	za92h06.s1 Homo sapiens cDNA clone 300059 3'.
138	CATGGTGGTGGACAC	H765573	1	4	13	6	13	H90469	yv01e06.r1 Homo sapiens cDNA clone 241474 5' simil
								R76765	yi63g01.r1 Homo sapiens cDNA clone 143952 5' simil
								T35045	EST79335 Homo sapiens cDNA similar to None..
139	CATGTGGGGTACCTT	H961304	0	6	13	2	9	H51447	yo31a05.r1 Homo sapiens cDNA clone 179504 5'.
								W46469	zc32c05.r1 Soares senescent fibroblasts NbHSF Homo
								W51800	zc48e04.r1 Soares senescent fibroblasts NbHSF Homo
								R33196	yh77f08.r1 Homo sapiens cDNA clone 135783 5'.
140	CATGTTCATTATAAT	H1003313	1	10	13	8	10	J04799	Human prothymosin-alpha
141	CATGCTTCTGTGTAC(T)	H515821	0	5	13	8	12	D80012	Human KIAA0190 protein
142	CATGACTGGCGAAGT	H125315	1	5	13	2	5	U02389	Human hLON ATP-dependent protease mRNA
								T29819	EST96617 Homo sapiens cDNA 5' end similar to ATP-d
143	CATGGAAAGAGCTGA	H526495	1	3	13	1	6	X14850	Human histone H2A.X.
144	CATGCAACTCTATGG	H269775	0	1	13	1	2	J04088	Human DNA topoisomerase II (top2) mRNA
145	CATGAAATTTGGTGC	H16303	0	0	13	0	0	K01891	Human beta globin retrovirus-like repetitive element
								H88396	EST28e05 Homo sapiens cDNA clone 28e05
146	CATGCTGCACTTACT	H496114	1	2	13	1	8	X74796	H.sapiens p85Mcm mRNA.
								D28480	Human mRNA for hMCM2, complete cds.
								D55716	Human B lymphoma mRNA for P1cdc47, complete cds.
147	CATGAATATTGAGAA	H53129	0	5	13	6	11	T30327	EST14849 Homo sapiens cDNA 5' end similar to None.
								T34394	EST66942 Homo sapiens cDNA 5' end similar to None.
								T47475	yb14c03.r1 Homo sapiens cDNA clone 71140 5'.
								T50289	yb14h08.r1 Homo sapiens cDNA clone 71199 5'.
148	CATGTCGCCGGGCGC	H890535	0	1	13	2	1		Unknown
149	CATGGGGGCAGCCG	H697495	0	2	13	2	7	H59914	Unknown
150	CATGCCAAGAAAGAA	H329737	0	6	12	4	4	U33818	Human inducible poly(A)-binding protein
151	CATGTTTTTGATAAA	H1048113	0	5	12	4	12	D16891	Human HepG2 3' region cDNA, clone hmd2c11.
152	CATGTGTGGAGAGCC	H977034	0	0	12	0	0	M29882	Human apolipoprotein A-II
153	CATGCCCACGGTTAG	H345789	0	5	12	5	4	Z49216	H.sapiens mitoxantrone-resistance associated mRNA.
154	CATGAATTCTCCTAA	H63325	0	1	12	1	1		Unknown
155	CATGGACCTCCGGGC	H548203	0	0	12	0	0		Unknown
156	CATGTGAATCTGGGT	H921067	0	2	11	7	8	M93651	Human set gene

157	CATGTCCTTCTCCAC	H884181	0	5	11	14	8	X15804	Human alpha-actinin.
158	CATGTATCTGTCTAC	H843485	0	4	11	2	3	T19569	609F Homo sapiens cDNA clone 609 similar to SET protein
159	CATGACGTTCTCTTC	H114144	0	0	11	1	17	Z36249	HHEA18W H. sapiens partial cDNA sequence; clone HEA18W;
160	CATGCCCTGAGTCAG	H358581	0	0	11	0	0	AA207189	zq73e07.r1 Stratagene neuroepithelium (#937231)Homo sapiens cDNA clone 647268 5' similar to TR:E16910 E16910 ENDONUCLEASE. ;
161	CATGGAATTCCTCGA	H540023	0	3	11	3	1	JN80776	za98h04.s1 Homo sapiens cDNA clone 300631 3'.
								AA025809	ze90d01.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366241 3'
								AA279492	zs85h05.s1 Soares NbHTGBC Homo sapiens cDNA clone 704313 3'
162	CATGGACGCCGAACT	H550274	0	1	11	6	0		Unknown
163	CATGGCGGACTGGGG	H631275	0	0	11	1	0	AA098867	zk84f04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 489535 3' similar to SW:A5_XENLA P28824 A5 PROTEIN PRECURSOR
164	CATGGGAACACACAG	H656453	0	1	11	0	2	R48460	yj67c12.r1 Homo sapiens cDNA clone 153814 5'.
								AA173819	zp01c02.r1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 595106 5'
165	CATGTTGCGGAGCCC	H1022502	0	2	11	2	1	L19183	HUMMAC30X Human MAC30 mRNA, 3' end.
								H61710	yr24a07.s1 Homo sapiens cDNA clone 206196 3'.
								H77330	yu11f12.s1 Homo sapiens cDNA clone 233519 3'.
								N69482	za18d05.s1 Homo sapiens cDNA clone 292905 3'.
166	CATGGCAGACATTGA	H598335	0	7	10	4	9	H41078	yp52c11.s1 Homo sapiens cDNA clone 191060 3' simil
167	CATGCACTTGAAAA	H294401	0	1	10	5	0	H04630	yj49g03.r1 Homo sapiens cDNA clone 152116 5'.
168	CATGGGTTGGCAGG	H719435	0	0	10	24	0	R77027	yi66e12.r1 Homo sapiens cDNA clone 144238 5'.
169	CATGTTCCTCGGGC	H1007018	0	1	10	4	12	R32331	yh68g02.s1 Homo sapiens cDNA clone 134930 3' simil
170	CATGCTGCCGAGCT	-497192	0	8	10	1	10	T86566	yd77g07.r1 Homo sapiens cDNA clone 114300 5' simil
171	CATGGTGAAAAAAA	H753665	0	2	10	3	7	S77357	transcript ch111 [human, RF1,RF48 stomach cancer c
172	CATGCTGTGCAGCA	H506149	0	6	10	6	1	M34338	Human spermidine synthase
173	CATGTAGTTTGTGG	-835515	0	1	10	0	2	U03911	Human mutator gene (hMSH2)
174	CATGATGTAGTAGTG	H242380	0	5	10	9	7	D55671	Human heterogeneous nuclear ribonucleoprotein
175	CATGGACCCACTACC	H545906	0	1	10	3	1	J03569	Human lymphocyte activation antigen 4F2 large subunit
176	CATGAAATAGGTTTT	H12992	0	1	10	6	3	D53402	Human fetal brain cDNA 5'-end GEN-108D03.
								T61971	yb96f02.r1 Homo sapiens cDNA clone 79035 5'.
								D61243	Human fetal brain cDNA 5'-end GEN-171G06.
								N77240	yv44d02.r1 Homo sapiens cDNA clone 245571 5'.
177	CATGCCGGGCGTGGT	H371131	0	0	10	1	2	T35761	EST90898 Homo sapiens cDNA 5' end similar to EST c

178	CATGGACTGAGCTTG	H555168	0	8	10	3	3	T31901	EST40719 Homo sapiens cDNA 5' end similar to None.
179	CATGAAACGCCCAAT	H6481	0	2	10	1	3	X98264	HSMPP41 H.sapiens mRNA for M-phase phosphoprotein, mpp4, 1523bp
180	CATGATGAGGCCGGG	H232027	0	4	10	7	1		Unknown
181	CATGGCCCACATCCG(A)	H610614	0	9	10	6	2	D87433	Human mRNA for KIAA0246 gene, partial cds

Table 3 - Transcripts decreased in colon cancer

Transcripts decreased in only colon primary tumors compared to normal colon (51 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag sequence	Tag_Number	NC	CT	CL	PT	PC	Accession	Gene Name
1	CATGGCTTTATTTGT	H654591	184	110	185	203	111	X00351	Human mRNA for beta-actin.
2	CATGCTAGCCTCACG	H468434	170	61	130	80	75	X04098	Human mRNA for cytoskeletal gamma-actin.
3	CATGCAAACCATCCA	H263478	137	83	245	36	502	X12883	Human mRNA for cytokeratin 18.
4	CATGCTTCCAGCTAA	H513181	64	23	36	53	104	D00017	Human lipocortin II mRNA.
5	CATGCCCCAGTTGCT	H348922	61	27	38	37	46	X04106	Human mRNA for calcium dependent protease (small subunit)
6	CATGGATGACCCCCC	H581974	53	4	42	6	32	Z65513	H.sapiens CpG island DNA genomic Mse1 fragment, cl
7	CATGCTGTACAGACA	H504098	50	22	26	6	32	W61077	zd30d02.r1 Soares fetal heart NbHH19W Homo sapiens
8	CATGCGGACTCACTG	H427848	47	15	26	18	4	D60944	Human fetal brain cDNA 5'-end GEN-141D02.
9	CATGCCCCCGCGGAA	H349801	47	10	21	15	8		Unknown
10	CATGCCTGGAAGAGG	H387107	46	19	39	47	14	J02783	Human thyroid hormone binding protein (p55) mRNA,
11	CATGGCCTGGCCATC	H621140	46	19	24	16	20	N33042	yy05d05.s1 Homo sapiens cDNA clone 270345 3'
12	CATGAGCAGGAGCAG	H150053	43	12	26	24	20	W07627	zb06a05.r1 Soares fetal lung NbHL19W Homo sapiens
13	CATGAACGTGCAGGG	H28235	42	6	57	2	10	X01630	Human mRNA for argininosuccinate synthetase.
14	CATGGCCGCCCTGCA	H615802	40	12	16	17	8	D43682	Human mRNA for very-long-chain acyl-CoA dehydrogen
15	CATGTGGGGAGAGGA	H960651	40	5	36	10	5	D29146	Human keratinocyte cDNA, clone 173.
16	CATGGCTGCCCTTGA	H648575	38	10	20	6	39	K00557	human alpha-tubulin mRNA, 3' end.
17	CATGTGGCCATCTGC	H955615	37	5	15	19	18	AA341633	AA341633 EST47188 Fetal kidney II Homo sapiens cDNA 5' end
18	CATGCGTTCCTGCGG	H456167	35	4	36	8	0	X77956	H.sapiens Id1 mRNA.
19	CATGTGCATCTGGTG	H937452	33	9	14	13	10	X87949	H.sapiens mRNA for BiP protein.
20	CATGGTGACCTCCTT	H755160	33	7	12	6	31	J04823	Human cytochrome c oxidase subunit VIII (COX8) mRNa
21	CATGTAGCTCTATGG	H826831	33	5	18	9	13	U16798	Human Na,K-ATPase alpha-1 subunit mRNA, complete c
22	CATGGTGCGCTAGGG	H760267	29	7	26	19	27	R50350	gb|R50350|R50350 yj59c04.s1 Homo sapiens cDNA clone 153030 3'.
								R50013	yj59c04.r1 Homo sapiens cDNA clone 153030 5'.
								C02981	Human Heart cDNA, clone 3NHC0642.

23	CATGGGGCGCTGTGG	H694767	28	6	20	6	26	T31329	EST30445 Homo sapiens cDNA 5' end similar to ubiquinol cytochrome-c reductase, 6.4 kDa.
24	CATGCCTCCAGTAC	H382130	27	6	12	3	19		Unknown
25	CATGCCTGTGACAGC	H388627	27	3	14	8	7	H63643	yr34d11.r1 Homo sapiens cDNA clone 207189 5' simil
26	CATGTCACAGTGCCT	H856806	24	5	8	17	11	W60924	zd27c08.r1 Soares fetal heart NbHH19W Homo sapiens
27	CATGAATAAAGGCTA	H49320	23	5	7	11	13	L25081	Human GTPase (rhoC) mRNA, complete cds.
28	CATGTTGTTGTTGAA	H1031929	23	5	13	15	25	D45887	Human mRNA for calmodulin, complete cds.
29	CATGAAGGTAGCAGA	H44179	23	4	10	16	12	N62815	yy66b11.s1 Homo sapiens cDNA clone 278493 3'.
30	CATGGTGTTGGGGGT	H769707	21	2	5	14	10	R68653	yi14b06.s1 Homo sapiens cDNA clone 139187 3'.
31	CATGTGCAGCGCCTG	H936344	21	1	5	7	13	X90858	H.sapiens mRNA for uridine phosphorylase.
32	CATGATGGCACGGAG	H238697	20	2	4	0	3	H19458	yn54c02.s1 Homo sapiens cDNA clone 172226 3' simil
33	CATGGCCAGACACCC	H608326	20	1	6	1	9	T30468	EST17149 Homo sapiens cDNA 5' end similar to None.
34	CATGCTTCTTGCCCC	H515990	20	0	17	3	0	V00491	Human gene for alpha 1 globin.
35	CATGACCCACGTCAG	H86453	19	2	7	22	9	X51345	Human jun-B mRNA for JUN-B protein.
36	CATGGGCTGCCTGCC	H686458	18	3	4	5	8	R72429	yj90e08.s1 Homo sapiens cDNA clone 156038 3'.
								R48449	yj67b10.s1 Homo sapiens cDNA clone 153787 3'.
								R52128	yj72b03.s1 Homo sapiens cDNA clone 154253 3'.
37	CATGGAGGGCCGGTG	H567660	18	2	14	6	16	X12910	Human Na+,K+ ATPase gene exons 1 - 3 (alpha III is
38	CATGGATGAATCCGG	H581847	17	1	3	2	2		Unknown
39	CATGAGCCCGACCAC	H153109	16	2	11	7	5	X81006	H.sapiens HCG I mRNA.
40	CATGGTTCAGCTGTC	H774780	16	2	12	3	12	L08666	Homo sapiens porin (por) mRNA, complete cds and tr
41	CATGCCTCGCTCAGT	H383443	16	1	8	6	7	U04627	Human 78 kDa gastrin-binding protein mRNA, complet
42	CATGCAAATAAAAGT	H265219	15	1	8	9	0	U17077	Human BENE mRNA, partial cds.
43	CATGTGCCGCCCGCA	H940378	15	1	8	0	3	U28369	Human semaphorin V mRNA, complete cds.
44	CATGGCAGTGGCCTC	H601752	15	0	6	4	3	D12038	Human HepG2 3'-directed MboI cDNA, clone s150.
45	CATGCTGGGCCTGAA	H502137	14	0	3	3	18	U77396	Human TNF-alpha inducible responsive element mRNA,
46	CATGGCCCATTGGAG	H611305	13	1	6	13	17	Z29093	H.sapiens EDDR1 gene for receptor tyrosine kinase.
47	CATGAAGAAAACCTC	H32792	12	0	2	2	0	T94990	ye38a04.s1 Homo sapiens cDNA clone 119982 3'.
								N69310	za25g05.s1 Homo sapiens cDNA clone 293624 3'.
								N98502	zb86e03.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 310492 3'
48	CATGGAATGATTTCT	H538878	12	0	6	6	14	F18838	H.sapiens EST sequence (007-X1-01) from skeletal m
49	CATGGCCTGGTCCTT	H621272	12	0	3	3	8	AA226928	zr21b10.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 664027 3'
50	CCATGGCCCACACAG	H610579	11	0	1	1	0	M60047	Human heparin binding protein (HBp17) mRNA

51	CATGGGATTCCAGTT	H671052	11	0	4	3	2	W52456	zc45e09.r1 Soares senescent fibroblasts NbHSF Homo

Transcripts decreased in both colon primary tumors and colon cancer cell lines compared to normal colon (130 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag Sequence	Tag_Number	NC	TU	CL	PT	PC	Accession	Gene Name
1	CATGCCTCCAGCTAC	H382109	803	191	304	136	663	X12882	Human mRNA for cytokeratin 8.
2	CATGCTAAGACTTCA	H460926	708	282	402	142	497	F15636	H.sapiens mitochondrial EST sequence (002T15)
3	CATGGCCCAGGTCAC	H610997	705	58	2	2	1		Unknown
4	CATGACCCTTGGCCA	H90022	512	348	93	43	235	F16940	H.sapiens mitochondrial EST sequence (009-T1-21) f
5	CATGACATTGGGTGA	H81583	504	92	4	0	0	M10050	Human liver fatty acid binding protein (FABP) mRNA
6	CATGGCGAAACCCTG	H622680	486	108	27	30	13	S61953	c-erbB3=receptor tyrosine kinase {alternatively sp
7	CATGAGCCCTACAAA	H153361	367	242	132	71	204	F15506	H.sapiens mitochondrial EST sequence (1-t-02) from
8	CATGGACCCAAGATA	H545828	276	131	0	7	0	T39321	ya04c01.r2 Homo sapiens cDNA clone 60480 5'.
								H24673	yl41a01.s1 Homo sapiens cDNA clone 160776 3'.
								D25586	HUMGS02706 Human colon 3'directed MboI cDNA, HUMGS02706, clone cm1673.
								T96160	ye09b02.s1 Homo sapiens cDNA clone 117195 3'.
9	CATGGCCGGGTGGGC	H617195	256	88	148	144	178	X64364	H.sapiens mRNA for M6 antigen.
10	CATGTTGGGGTTTCC	H1026814	202	75	84	235	369	M11146	Human ferritin H chain mRNA, complete cds.
11	CATGCTCCACCCGAA (or G)	H479577	201	120	0	11	3	L15203	Human secretory protein (P1.B) mRNA, complete cds.
12	CATGGCAGGGCCTCA	H600670	196	68	6	32	19	X93036	H.sapiens mRNA for MAT8 protein.
13	CATGATCGTGGCGGG	H224923	194	24	97	40	39	H93844	yv07h09.r1 Homo sapiens cDNA clone 242081 5' similar to SP:A39484 A39484 ANDROGEN-WITHDRAWAL APOPTOSIS PROTEIN RVP1,
14	CATGCAAGCATCCCC	H271574	190	99	101	30	139	F17001	H.sapiens mitochondrial EST sequence (011-T1-13) f
15	CATGGACATCAAGTC	H544012	189	33	76	57	219	Y00503	Human mRNA for keratin 19.
16	CATGGTTGTGGTTAA	H782013	178	110	14	340	139	W16632	zb05a11.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 301148 5' similar to gb:V00567 BETA-2-MICROGLOBULIN PRECURSOR (HUMAN);.
								AA143804	zo31h04.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588535 3'

								AA133597	97 zl92h02.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 512115 3
								T53199	ya86c05.s1 Homo sapiens cDNA clone 68552 3'.
								R00081	ye73c04.s1 Homo sapiens cDNA clone 123366 3'.
17	CTAGTGCTCCTACCC	H947654	174	27	1	0	0	M16364	Human creatine kinase-B mRNA, complete cds.
18	CATGCACCCCTGATG	H284132	172	33	26	3	6	R09410	yf22e12.s1 Homo sapiens cDNA clone 127630 3' similar to contains Alu repetitive element
19	CATGCCGCTGCACTC	H368200	163	40	4	10	4	C01918	HUMGS0003915, Human Gene Signature, 3'-directed cDNA sequence.
								R92735	yq04h09.s1 Homo sapiens cDNA clone 196001 3' similar to contains Alu repetitive element
								W90374	zh78e12.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 418222 3' similar to contains Alu repetitive element
								X52003	H.sapiens pS2 protein gene.
20	CATGCTGGCCCTCGG	H501111	163	20	0	26	1	M18981	Human prolactin receptor-associated protein (PRA)
21	CATGCCCCCTGGATC	H350116	160	40	24	88	181	M64303	Human galactoside-binding protein mRNA.
22	CATGTTCACTGTGAG	H1001401	160	34	13	74	71	X16455	Human mRNA for carcinoembryonic antigen pCEA80-11.
23	CATGATTGGAGTGCT	H256186	155	34	1	11	6	U14943	Human MHC antigen (HLA-B) mRNA, complete cds.
24	CATGCTGACCTGTGT	H493039	149	44	32	98	37	M81457	Human calpactin I light chain mRNA, complete cds.
25	CATGAGCAGATCAGG	H149715	145	50	88	156	130	C21047	HUMGS0002546, Human Gene Signature, 3'-directed cDNA sequence
26	CATGGGAAAACAGAA	H655433	126	37	0	24	16	AA132779	zo21h08.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 587583 3' similar to SW:LEG4_RAT P38552 GALECTIN-4
								AA054072	zl68h06.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 509819 3'
								AA132736	zo18g08.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 587294 3' similar to SW:LEG4_RAT P38552 GALECTIN-4
								X04412	Human mRNA for plasma gelsolin.
27	CATGTCACCGGTCAG	H857781	122	7	7	30	7	X77658	H. sapiens mRNA for HLA-B*7301.
28	CATGTGCAGCACGAG	H936217	122	26	32	84	2	AA146606	zo35c09.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588880 3'
29	CATGGGAACTGTGAA	H657337	115	7	1	14	21	AA146775	zo35g09.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588928 3'
								AA161043	zo74g11.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592676 3'

								AA088704	zl83f08.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511239 3'
30	CATGCGAGGGGCCAG	H404117	114	32	54	60	40	H00427	yj23g11.r1 Homo sapiens cDNA clone 149636 5'.
								AA158715	zo63d03.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 591557 3'
								T08562	EST06454 Homo sapiens cDNA clone HIBBG31 3' end.
								AA078845	zm21a12.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526270 3'
31	CATGTAAATTGCAAA	H790417	113	6	1	0	0	X73502	H. Sapiens mRNA for cytokeratin 20.
32	CATGGGCTGGGGGCC	H686762	113	36	48	45	43	J03191	Human profilin mRNA, complete cds.
33	CATGGTGCTGAATGG	H761359	109	20	30	67	111	U02629	Human smooth muscle myosin alkali light chain mRNA
34	CATGGTGCACTGAGC	H758243	107	13	36	34	82	X07059	Human M4-50 mRNA for HLA class I antigen.
35	CATGTTTAACGGCCG	H1032614	107	31	14	3	37	F15592	H.sapiens mitochondrial EST sequence (001T24) from
36	CATGCCCTCCCGAAG	H357729	106	17	7	3	6	AA053660	zl74e07.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 510372 3' similar to contains Alu repetitive element
								D25711	HUMGS04077 Human colon 3'directed MboI cDNA, HUMGS04077, clone cm1210
37	CATGAGGTGGCAAGA	H178755	105	15	22	14	27	Z56800	H.sapiens CpG DNA, clone 140c4, reverse read cpg14(Mitochondria EST
38	CATGATACTCCACTC	H204104	102	11	0	0	0	M95174	Human guanylin mRNA, complete cds.
39	CATGCTCGCGCTGGG	H484987	101	25	5	4	16		Unknown
40	CATGGGGGCAGGGCC	H697514	82	32	28	37	65	R90863	yn01b01.r1 Homo sapiens cDNA clone 167113 5' similar to SP:ZK783.1 CE00760 ;.
								T24702	EST277 Homo sapiens cDNA clone 10H4.
41	CATGGAAGCAGGACC	H533666	80	33	42	28	87	X95404	H.sapiens mRNA for non-muscle type cofilin.
42	CATGCCAGGGGAGAA	H338569	75	22	28	30	16	X67325	H.sapiens p27 mRNA.
43	CATGACACAGCAAGA	H70211	74	31	30	10	31	F16604	H.sapiens mitochondrial EST sequence (009T28) from
44	CATGAGAATAGCTTG	H134304	69	29	1	3	0	N69361	za16a03.s1 Homo sapiens cDNA clone 292684 3' similar to contains Alu repetitive element;contains element L1 repetitive element
								AA015918	ze30b10.s1 Soares retina N2b4HR Homo sapiens cDNA clone 360475 3' similar to contains Alu repetitive element
								H26689	yl14h01.s1 Homo sapiens cDNA clone 158257 3' similar to contains Alu repetitive element;contains TAR1 repetitive element ;.
45	CATGCGCTGTGGGGT	H424875	68	9	6	5	23	AA256365	zr79h11.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 681957 3' similar to WP:C33A12.7 CE05353

								W47357	zc39e11.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 324716 3'
								W19276	zb90f03.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 310877 3'
								R07159	yf13h12.s1 Homo sapiens cDNA clone 126791 3'.
46	CATGCATAGGTTTAG	H314109	68	5	0	0	0	L02785	Homo sapiens colon mucosa-associated (DRA) mRNA
47	CATGGCCGACCAGGT	H614731	65	19	0	3	6	U11862	Human clone HP-DAO1 diamine oxidase
48	CATGAGCTCTTGGAG	H161769	64	11	1	1	2	N93240	zb68b06.s1 Homo sapiens cDNA clone 308723 3'.
								T16906	NIB1986 Normalized infant brain, Bento Soares Homo sapiens cDNA 3'end.
								H78256	yu22h07.s1 Homo sapiens cDNA clone 234589 3' similar to SP:SBP_MOUSE P17563 SELENIUM-BINDING
								T32362	EST47523 Homo sapiens cDNA 3' end similar to similar to Selenium-binding protein,liver.
49	CATGCCCAACGCGCT	H344474	57	1	0	3	0	V00493	Human messenger RNA for alpha globin.
50	CATGGACGCGGCGCG	H550554	55	21	2	7	14		Unknown
51	CATGACCCCCCCGCC	H87386	54	16	15	15	3	X51346	Human jun-D mRNA for JUN-D protein.
52	CATGATGCGGGAGAA	H236169	52	6	10	11	7	R34039	yh83f04.r1 Homo sapiens cDNA clone 136351 5'.
								H03961	yj44e07.s1 Homo sapiens cDNA clone 151620 3'.
								R33498	yh83f04.s1 Homo sapiens cDNA clone 136351 3'.
53	CATGTCAGCTGCAAC	H862097	51	6	0	0	0	AA053043	zl71e06.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 510082 5'
54	CATGGTAAGTGTACT	H723890	50	14	15	1	30	F17394	H.sapiens mitochondrial EST sequence (007T13) from
55	CATGTGTGGGTGCTG	H977640	49	20	17	21	8	Z13009	H.sapiens mRNA for E-cadherin.
56	CATGGCTGTGCCTGG	H650847	48	17	15	8	31	X15505	Human mRNA for pancreatic trypsinogen III.
57	CATGTGAGTGACAGA	H929299	48	4	0	0	0	H14641	yl26g02.s1 Homo sapiens cDNA clone 159410 3'.
58	CATGGGCTGGGCCTG	H686744	47	11	13	32	8	M20469	Human brain-type clathrin light-chain b mRNA,
59	CATGTAATCCCAGCA	H800074	46	15	5	8	11	N50873	yy92c07.s1 Homo sapiens cDNA clone 281004 3' similar to contains Alu repetitive element;contains element MER32 repetitive element
60	CATGGACCAGTGGCT	H545514	45	1	0	0	1	U79725	Human A33 antigen precursor mRNA, complete cds
61	CATGGGCACCGTGCT	H673210	44	10	1	14	14		Unknown
62	CATGAAGGACCTTTT	H41344	43	17	14	22	24	H11216	ym14f06.r1 Homo sapiens cDNA clone 47991 5'.
								H52178	yt85h08.s1 Homo sapiens cDNA clone 231135 3'.
								T40539	ya05b02.s1 Homo sapiens cDNA clone 60555 3'.

								AA303091	EST12940 Uterus tumor I Homo sapiens cDNA 3' end
63	CATGGCAGCTCCTGT	H599903	43	8	17	24	13	W02429	za52d02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 296163 5'.
								N20325	yx44c11.s1 Homo sapiens cDNA clone 264596 3'.
								N45127	yz13c12.s1 Homo sapiens cDNA clone 282934 3'.
								N90407	zb38c11.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 305876 3'.
64	CATGTGTCCTGGTTC	H972720	43	12	14	25	5	U03106	Human wild-type p53 activated fragment-1 (WAF1) mR
65	CATGACAAACCCCCA	H65878	42	16	7	12	11	W37827	zc11f01.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322009 3'
								W15332	gb|W15332|W15332 zc16d10.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 322483 3'
								W32410	zc04g10.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 321378 3'
								N32312	yw82c01.s1 Homo sapiens cDNA clone 258720 3'.
66	CATGTAGGATGGGGG	H828331	41	6	11	6	9	U51478	Human sodium/potassium-transporting ATPase beta-3
67	CATGACTGTGGCGGC	H126619	41	7	1	4	35		Unknown
68	CATGGTAGCAGGTGT	H730287	40	7	13	17	24	AA180815	zp44f11.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 612333 3' similar to contains Alu repetitive element;
								R34696	yh87e04.s1 Homo sapiens cDNA clone 136734 3' similar to contains Alu repetitive element;.
								R34696	yh87e04.s1 Homo sapiens cDNA clone 136734 3' similar to contains Alu repetitive element;.
								AA194497	zq06e03.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 628924 3' similar to contains Alu repetitive element
69	CATGAATCACAAATA	H53508	40	12	0	3	0	T11144	hbc760 Homo sapiens cDNA clone hbc760 3'end similar to nonspacific crossreacting antigen.
								AA058357	zl67e01.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 509688 3' similar to TR:G189087
								C05803	similar to none
70	CATGAGGATGGTCCC	H167606	40	11	4	4	5	AA143765	zo31e02.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588506 3'
								AA179299	zp45b09.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 612377 3'

71	CATGCCAAAGCTATA	H328308	38	11	6	2	18	M35252	Human CO-029.
72	CATGCGGGAGTCGGG	H434907	38	8	6	0	0	R87448	ym89c10.s1 Homo sapiens cDNA clone 166098 3'.
73	CATGGCCGTGGAGAG	H618121	38	9	5	17	26	X79882	H.sapiens lrp mRNA.
74	CATGCCCCCGAAGCC	H349706	37	6	0	0	0		Unknown
75	CATGATTTCAAGATG	H259108	37	1	0	0	0	J03037	Human carbonic anhydrase II mRNA, complete cds.
76	CATGGCCCAGTGGCT	H611050	37	3	0	2	10		Unknown
77	CATGATGGTGGGGGA	H241323	36	2	6	25	2	M92843	H.sapiens zinc finger transcriptional regulator mRNA
78	CATGCCTGCCCCCCT	H386390	35	12	7	7	5	X60188	Human ERK1 mRNA for protein serine/threonine kinase
79	CTAGTGGAAAGTGAA	H950457	34	1	1	12	0	V01512	Human cellular oncogene c-fos (complete sequence).
80	CATGGTCATCACCAC	H740629	34	0	0	0	0	U34279	Human uroguanylin mRNA, complete cds.
81	CATGCTTATGGTCCC	H511670	34	1	0	3	1	AA287021	zs57c03.s1 Soares NbHTGBC Homo sapiens cDNA clone 701572 3'
82	CATGCTGGGCCTCTG	H502136	34	3	4	11	5	T55226	yb47a01.s1 Homo sapiens cDNA clone 74280 3' containing L1 repetitive element
								R37446	yf56e10.s1 Homo sapiens cDNA clone 26129 3' similar to gb:X07173 INTER-ALPHA-TRYPSIN INHIBITOR COMPLEX COMPONENT II
								AA406180	zu65c08.s1 Soares testis NHT Homo sapiens cDNA clone 742862 3'
83	CATGGCCCAGGGCCC	H610982	33	3	0	0	2	R09752	Unknown
84	CATGTTTTTACTGAT	H1047673	33	7	0	4	2	R81530	yj02b10.r1 Homo sapiens cDNA clone 147547 5'.
								T32348	EST47211 Homo sapiens cDNA 3' end similar to None..
								W57810	zd17g02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 340946 3'
								AA398527	zt47e12.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 725518 3'
85	CATGCCTGCTTGTCG	H387054	32	2	1	6	32	X63187	H.sapiens HE4 mRNA for extracellular proteinase inhibitor homologue
86	CATGACCTGGGGAGG	H96931	32	6	4	8	6		Unknown
87	CATGCCTTCAAATCA	H390158	31	1	0	0	0	R46266	yg52g07.s1 Homo sapiens cDNA clone 36232 3' similar to gb:M33987 CARBONIC ANHYDRASE I
88	CATGTCGGAGCTGTT	H893564	30	1	4	7	1	H98618	yx12a06.s1 Homo sapiens cDNA clone 261490 3'.
								AA171705	zo97h01.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 594865 3'
								H99212	yx15g08.s1 Homo sapiens cDNA clone 261854 3'.

								AA029975	zk10e12.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470158 3'
89	CATGGGAGGTGGGGC	H666539	30	6	5	32	22	M75161	H.sapiens granulin mRNA, complete cds.
90	CATGTTCCACTAACC	H1003970	30	7	3	16	17	T30344	gb|U53204|HSU53204 Human plectin (PLEC1) mRNA, complete cds.
91	CATGGTCTGGGGGAT	H752297	29	1	3	9	3	T60135	yc22a06.s1 Homo sapiens cDNA clone 81394 3'.
								T30403	gb|U67963|HSU67963 Human lysophospholipase homolog (HU-K5) mRNA
92	CATGTTAACCCCTCC	H984414	29	5	0	18	0	R23595	yh39a12.r1 Homo sapiens cDNA clone 132094 5' similar to gb:D26129 RIBONUCLEASE PANCREATIC PRECURSOR (HUMAN)
								R69445	yj83c08.s1 Homo sapiens cDNA clone 155342 3' similar to gb:D26129 RIBONUCLEASE PANCREATIC PRECURSOR (HUMAN);.
								R79191	yi84h01.s1 Homo sapiens cDNA clone 145969 3' similar to gb:D26129 RIBONUCLEASE PANCREATIC PRECURSOR (HUMAN);.
								R49965	yj56c03.s1 Homo sapiens cDNA clone 152740 3' similar to gb:D26129 RIBONUCLEASE PANCREATIC PRECURSOR (HUMAN);.
93	CATGATGACGCTCAC	H231029	28	5	5	4	6	AA410947	zv35h12.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 755687 5' similar to TR:G459890 G459890 OVEREXPRESSED IN TESTICULAR TUMORS
								H02520	yj40c11.r1 Homo sapiens cDNA clone 151220 5'.
								AA130551	zo12g08.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 586718 5' similar to TR:G459890 G459890 OVEREXPRESSED IN TESTICULAR TUMORS.
94	CATGCACCTGTCATC	H286420	28	5	0	5	4	W68230	zd33c10.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 342450 3' similar to contains Alu repetitive element
								R89822	yp90a02.s1 Homo sapiens cDNA clone 194666 3' similar to contains Alu repetitive element;
								AA053322	zk69e08.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 488102 3' similar to contains element MER6 repetitive element
95	CATGGATCCCAACTG	H578824	27	1	1	24	17	V00594	Human mRNA for metallothionein from cadmium-treated cells
96	CATGCTTAGAGGGGT	H510123	27	1	5	9	6	H43742	yp21d05.r1 Homo sapiens cDNA clone 188073 5' similar to gb:J05021 EZRIN
97	CATGATGGCCCATAC	H238925	27	4	3	1	0		emb|Y09616|HSICE H.sapiens mRNA for putative carboxylesterase
98	CATGGCAAGAAAGTG	H591884	27	1	0	2	0	V00497	Human messenger RNA for beta-globin.

99	CATGTACCTCTGATT	H810468	27	5	7	11	12	X65614	H.sapiens mRNA for calcium-binding protein S100P.
100	CATGATGATGGCACC	H233106	26	0	2	0	2		
101	CATGTTCTGTAGCCC	H1014566	25	5	0	4	0		emb|Z69881|HSSERCA3M H.sapiens mRNA for adenosine triphosphatase, calcium
102	CATGCCTGTCTGCCA	H388582	24	1	2	1	3	T99568	ye65c02.r1 Homo sapiens cDNA clone 122594 5'.
								T87539	yd89f09.s1 Homo sapiens cDNA clone 115433 3'.
103	CATGTATGATGAGCA	H844682	23	4	0	1	0		gb|AA347726|AA347726 EST54132 Fetal heart II Homo sapiens cDNA 5' end similar to transmembrane secretory component
104	CATGCTGGCAAAGGT	H500747	23	0	0	0	0		
105	CATGCTTGATTCCCA	H517078	23	4	4	17	7	L42379	Homo sapiens bone-derived growth factor (BPGF-1) m
106	CATGCTTGACATACC	H516402	22	0	0	7	2	X68277	H.sapiens CL 100 mRNA for protein tyrosine phosphase
107	CATGGCTGGCACATT	H649492	22	5	0	0	0	M82962	Human N-benzoyl-L-tyrosyl-p-amino-benzoic acid hydrolase alpha subunit (PPH alpha) mRNA, complete cds
108	CATGTCTGAATTATG	H909556	21	1	1	1	1	X16354	Human mRNA for transmembrane carcinoembryonic antigen (CEA)
109	CATGGGAAGAGCACT	H657554	21	1	1	3	3	X74570	H.sapiens mRNA for Gal-beta(1-3/1-4)GlcNAcalpha-2,3-sialyltransferase
110	CATGGCTCTTCCCCA	H646998	20	2	0	1	0	R87768	yo45d01.s1 Homo sapiens cDNA clone 180865 3' similar to contains PTR5 repetitive element
								R85880	yo36g07.s1 Homo sapiens cDNA clone 180060 3' similar to contains PTR5 repetitive element
111	CATGAAATCTGGCAC	H14245	20	2	0	4	3	L20826	Human I-plastin mRNA, complete cds.
112	CATGTAATTTGCATT	H802708	19	2	0	1	7	Z50751	HSB4BMR H.sapiens mRNA for B4B
								U77085	Human epithelial membrane protein (CL-20) mRNA, complete cds
								Y07909	HSPAPR H.sapiens mRNA for Progression Associated Protein
113	CATGGTGGGGGCGCC	H764570	18	1	1	8	2	R48529	yj64g10.r1 Homo sapiens cDNA clone 153570 5'.
114	CATGTTATGGTGTGA	H998127	17	0	0	1	0	T27534	EST10a24 Clontech adult human fat cell library HL1108A Homo sapiens cDNA clone 10a24.
115	CATGGGAGAAACAGC	H663571	17	1	2	4	0	T86124	yd84b04.s1 Homo sapiens cDNA clone 114895 3'.
								AA131008	zo15g05.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 587000 3'
								R49945	yj58g11.s1 Homo sapiens cDNA clone 152996 3'.
								T57044	ya84h01.s1 Homo sapiens cDNA clone 68401 3'.
116	CATGCCAACACCAGC	H328787	17	1	0	0	0		
117	CATGAGGTGACTGGG	H178299	17	0	0	0	0		
118	CATGGCCATCCTCCA	H609654	16	0	0	0	0		gb|R73013|R73013 yj94a09.r1 Homo sapiens cDNA clone 156376 5'.

119	CATGTTTCTCGTCGC	H1039799	15	1	0	4	4	M69013	Human guanine nucleotide-binding regulatory protein
120	CATGTCAGAGCGCTG	H860776	15	1	1	1	0		Unknown
121	CATGTTCCGCGTTCC	H1006014	14	1	0	0	2	N58523	yv72h06.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 248315 3' similar to contains element PTR7 repetitive element
122	CATGTACGGTGTGGG	H814011	14	1	0	0	0		Unknown
123	CATGCTCAGAACTTG	H477216	14	0	1	4	13		Unknown
124	CATGGGACTAAATGA	H662543	13	1	0	1	0	M29540	Human carcinoembryonic antigen mRNA (CEA), complete cds.
125	CATGGCTTGGGGATT	H653988	12	0	0	0	1	D25786	HUMGS04154 Human colon 3'directed MboI cDNA, HUMGS04154, clone cm0215.
								T73613	yc36e02.r1 Homo sapiens cDNA clone 82778 5' similar to gb:L07765 LIVER CARBOXYLESTERASE PRECURSOR
126	CATGACCCAACTGCC	H86138	12	0	0	0	1		Unknown
127	CATGCTGAACCTCCC	H491894	12	0	0	2	2		gb|T95615|T95615 ye40e03.s1 Homo sapiens cDNA clone 120220 3'.
128	CATGCAAGAGTTTCT	H271102	11	0	0	2	0	AA226797	zr19b11.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 663837 3'
								AA218730	zq97h01.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 649969 3'
129	CATGGTCCGAGTGCA	H743610	11	0	0	8	5	H38178	yp57f10.r1 Homo sapiens cDNA clone 191563 5' similar to gb:M90657 TUMOR-ASSOCIATED ANTIGEN L6 (HUMAN);.
130	CATGTTTGGTTTCAC	H1043445	11	0	0	0	0		Unknown

Transcripts decreased in only colon cancer cell lines compared to normal colon (78 genes)

NC: Normal Colon
TU: Colon Primary Tumor
CL: Colon Cancer Cell Line
PT: Pancreatic Primary Tumor
PC: Pancreatic Cancer Cell Line

#	Tag sequence	Tag_Number	NC	TU	CL	PT	PC	Accession	Gene Name
1	CATGCACCTAATTGG	H285759	612	755	411	161	333	F15516	H.sapiens mitochondrial EST sequence (1-t-12)
2	CATGATTTGAGAAGC	H260227	603	566	158	249	173	F12396	H. sapiens partial cDNA sequence; clone c-39e04.
3	CATGTGATTTCACTT	H933704	452	595	235	80	314	L08441	Human autonomously replicating sequence (ARS) mRNA
4	CATGTTCATACACCT	H1002566	444	357	114	64	191	F15553	H.sapiens mitochondrial EST sequence (001T14)
5	CATGCCACTGCACTC	H335432	385	402	223	278	132	X51525	Human cortex mRNA containing an Alu repetitive element
6	CATGACTAACACCCT	H114966	369	446	171	76	161	F16402	H.sapiens mitochondrial EST sequence (141-20)
7	CATGCACTACTCACC	H291282	293	527	78	14	83	U09500	Human mitochondrion cytochrome b gene, partial cds
8	CATGAAAACATTCTC	H1272	200	169	98	17	223	F15744	H.sapiens mitochondrial EST sequence (101-03)
9	CATGCTCATAAGGAA	H478249	184	127	70	21	75	F15511	H.sapiens mitochondrial EST sequence (1-t-07)
10	CATGTCGAAGCCCCC	H885334	147	183	94	49	57	F18587	H.sapiens mitochondrial EST sequence (022T19)
11	CATGACGCAGGGAGA	H103075	145	160	91	69	47	H03983	yj47a08.s1 Homo sapiens cDNA clone 151862 3'.
12	CATGTTGGCCAGGCT	H1025322	124	194	63	111	51	X74301	H.sapiens mRNA for MHC class II transactivator.
13	CATGTTGGTGAAGGA	H1027595	98	106	17	183	107	M17733	Human thymosin beta-4 mRNA, complete cds.
14	CATGATCACGCCCTC	H214616	97	186	17	41	49	U46913	Human EST overexpressed in pancreatic cancer (xs31)
15	CATGTGCCTGCACCA	H941638	67	48	25	75	34	X05607	Human mRNA for cysteine proteinase inhibitor precursor
16	CATGAGACCCACAAC	H136465	64	121	28	24	15	D54113	Human fetal brain cDNA 5'-end GEN-129B05.
17	CATGAGTTTGTTAGT	H196339	60	33	17	13	15	X14758	Human mRNA for adenocarcinoma-associated antigen
18	CATGGGAACAAACAG	H656389	56	41	4	31	3	L33930	Homo sapiens CD24 signal transducer mRNA
19	CATGTGGTGTATGCA	H965434	53	271	6	30	5	D50954	Human fetal brain cDNA 3'-end GEN-002A10.
20	CATGGAAATACAGTT	H527436	49	35	10	100	36	M11233	Human cathepsin D mRNA, complete cds.
21	CATGGTGGCTCACGC	H763719	49	37	21	27	15	U25801	Human Tax1 binding protein mRNA, partial cds.
22	CATGGTGGTGCACAC	H765509	45	26	18	23	15	U31215	Human metabotropic glutamate receptor 1 alpha
23	CATGGGGTTGGCTTG	H704160	44	56	2	6	1	S79597	tRNASer(UNC) {human, muscle, MERRF/MELAS overlap s
24	CATGGTGGCGGGTGC	H763567	42	32	15	20	5	T48809	yb05c03.r1 Homo sapiens cDNA clone 70276 5' contai
25	CATGTAGACTAGCAA	H821029	39	23	1	23	10	M69023	Human globin gene.

26	CATGGCTAGGTTTAT	H641789	38	144	13	25	13	D51017	Human fetal brain cDNA 3'-end GEN-007C04.
27	CATGGGCTTTAGGGA	H687915	37	372	6	29	11	W15552	zb91h11.s1 Soares parathyroid tumor NbHPA Homo sap
28	CATGGGGGTCAGGG	H699691	37	170	11	16	9	F16326	H.sapiens mitochondrial EST sequence (132-20) from skeletal muscle
29	CATGATTTTCTAAAA	H261569	33	13	11	8	2	AA315049	EST186995 HCC cell line (matastasis to liver in mouse) II Homo sapiens cDNA 5' end
30	CATGCACTTGCCCT	H294488	33	18	11	17	36	F01150	H. sapiens partial cDNA sequence; clone A6A03; ver
31	CATGCCTGCTGCAGG	H386963	32	13	0	6	2	N29971	yw53h01.s1 Homo sapiens cDNA clone 255985 3'.
32	CATGAGAACCTTCCA	H132598	32	14	3	16	12	K02883	Human MHC class I HLA-A2 gene, complete cds.
33	CATGCTCTGCCCTC	H489822	32	32	7	20	5	R09140	yf25f12.s1 Homo sapiens cDNA clone 127919 3'.
								R76005	yl22c10.s1 Homo sapiens cDNA clone 158994 3'.
								T33596	EST58371 Homo sapiens cDNA 3' end similar to None..
34	CATGGCCATCCCCTT	H609624	29	73	7	14	16	F16449	H.sapiens mitochondrial EST sequence (129-09)
35	CATGGCCCAGCGGCC	H610922	28	9	1	1	7	AA292959	zt54f10.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 726187 3'
36	CATGTGGCGCGTGTC	H956860	26	8	1	1	2	AA292466	zt31c11.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 723956 5' similar to TR:G205858 G205858 RAT ORF
								N92384	zb62d07.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 308173 3' similar to PIR:A39484 A39484 androgen-withdrawal apoptosis protein RVP1, prostatic - rat
								N80203	zb19c06.s1 Homo sapiens cDNA clone 302506 3' similar to PIR:A39484 A39484 androgen-withdrawal apoptosis protein RVP1, prostatic - rat ;
								AA039323	zk39d06.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 485195 3' similar to PIR:A39484 A39484 androgen-withdrawal apoptosis protein RVP1
37	CATGAGGGTGTTTTC	H175872	26	218	7	20	10	U21468	Human partial cDNA sequence with CCA repeat region
38	CATGCCTGGGAAGTG	H387596	25	10	0	45	17	M34088	Human episialin variant A mRNA, 3' end.
39	CATGAGTCTGCTGGA	H188027	24	9	1	0	0		Unknown
40	CATGCCCGCCTCTTC	H353760	24	11	2	3	4	T10098	seq816 Homo sapiens cDNA clone b4HB3MA-COT8-HAP-Ft
41	CATGAAAAGAGTGGT	H2235	22	9	2	0	7	X83228	H.sapiens mRNA for LI-cadherin.
42	CATGGCCACGTGGAG	H607977	21	7	1	2	2	L27415	Homo sapiens huntingtin (HD) gene, exon 66.
43	CATGAGGATGTGGG	H167659	21	5	4	1	3	C00470	dbj|C00470|C00470 HUMGS0007620, Human Gene Signature, 3'-directed cDNA sequence.
								N63531	yy62g08.s1 Homo sapiens cDNA clone 278174 3'.

								AA165679	zo80f04.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 593215 3'
44	CATGTATAGTCCTCT	H838494	20	7	1	3	4	AA411012	zv40a02.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 756074 3'
								AA133595	zl92g08.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 512126 3'
								AA292774	zt56b12.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 726335 3'
45	CATGGGTCCTCTCTT	H710520	20	7	2	2	2	R53216	yj73h02.r1 Homo sapiens cDNA clone 154419 5' simil
46	CATGATGGGCTTGAT	H240121	19	4	0	3	3	D20113	Human HL60 3'directed MboI cDNA, HUMGS01086, clone
47	CATGCTGCCCCCCAT	H496981	19	5	0	1	4		Unknown
48	CATGTTCTCTACACA	H1013522	19	4	1	8	2	U35048	Human TSC-22 protein mRNA, complete cds.
49	CATGAAGAAGCAGGG	H33355	18	4	2	2	8	R81767	yj05g03.r1 Homo sapiens cDNA clone 147892 5'.
50	CATGAGTAGGTGGCC	H183018	18	131	2	17	7	D51021	Human fetal brain cDNA 3'-end GEN-007D07.
51	CATGACAGTGTGTGT	H77551	18	5	3	0	8	D26146	Human DNA for putative protein kinase.
52	CATGGGAAAAGTGGT	H655547	18	13	3	70	1	M11465	Human alpha-1-antitrypsin mRNA, complete cds.
53	CATGAAGAAAGCTC	H32926	17	4	0	5	1	R78188	yi81g01.r1 Homo sapiens cDNA clone 145680 5'.
54	CATGACACCCATCAC	H70965	17	4	0	0	0	M22406	Human intestinal mucin mRNA, partial cds, clone SM
55	CATGAGATCCCAAGG	H144707	17	18	0	0	0	T24507	EST082 Homo sapiens cDNA clone 3E6..
								N79237	za63a11.s1 Homo sapiens cDNA clone 297212 3' similar to PIR:S49589 S49589 cortical granule lectin - African clawed frog ;.
								T31354	\EST30893 Homo sapiens cDNA 5' end similar to None..
56	CATGAATAGTTTCCC	H52214	16	4	0	0	0	H54696	yq92e02.s1 Homo sapiens cDNA clone 203258 3' simil
57	CATGCAGAAAGCATC	H295060	16	9	0	0	0	M22430	Human RASF-A PLA2 mRNA, complete cds.
58	CATGGCTTTGCTTTG	H654976	16	4	2	8	1	AA374631	EST86866 HSC172 cells I Homo sapiens cDNA 5' end
								AA137163	zn93g08.r1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone 565790 5'
								AA029320	zk10f05.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470145 3'
59	CATGTGCTGCATTGA	H948543	15	2	0	1	0	D25681	Human colon 3'directed MboI cDNA, HUMGS04047, clon
								AA253331	zr72g02.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 668978 3'
								H05110	yl75f07.s1 Homo sapiens cDNA clone 43778 3'.
60	CATGCCATCGTCCTT	H341720	15	8	1	1	10		Unknown
61	CATGGAACAGCTCAC	H529013	14	23	0	0	0	AA297150	EST112734 Colon I Homo sapiens cDNA 5' end

62	CATGGGGCTACGTCC	H695406	14	4	0	1	0	M25629	Human kallikrein mRNA, complete cds, clone clone p
63	CATGCCCGGCTCCTC	H354776	14	7	1	5	2	H18836	ym45d10.s1 Homo sapiens cDNA clone 51262 3'.
								AA026974	zk01e10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 469290 3'
								AA405031	zu12c12.r1 Soares testis NHT Homo sapiens cDNA clone 731638 5' similar to gb:M61900 Human prostaglandin D synthase gene, complete cds. (HUMAN);
64	CATGAGGTACTACTA	H176584	13	9	0	9	8	U66894	gb|U66894|HSU66894 Human epithelium-restricted Ets protein ESX mRNA,
								U73843	Human epithelial-specific transcription factor ESE-1b (ESE-1) mRNA, complete cds
65	CATGCAAATAAATTA	H265232	13	3	0	1	0	D25996	Human colon 3'directed MboI cDNA, HUMGS06772
66	CATGCTGTAAAAAAA	H503809	13	6	0	1	1		Unknown
67	CATGGTTCAATCCCT	H774358	13	3	0	2	0	AA071520	ze88g07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366108 3'
								N90742	za90h10.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 299875 3'.
								AA086292	zn52h06.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 561851 3'
68	CATGAATAAAGCCTT	H49304	12	4	0	0	0	D11499	Human HepG2 3'-directed MboI cDNA, clone a-35.
69	CATGGGAAGGTTTAC	H658173	12	2	0	1	0	T16031	IB2474 Homo sapiens cDNA 3'end.
70	CATGGGATGGCTTAT	H670333	12	1	0	6	1	T74426	yc82e01.r1 Homo sapiens cDNA clone 22306 5'.
71	CATGGGTGGCCCGGG	H715099	12	2	0	3	2	N73771	za61h02.s1 Homo sapiens cDNA clone 297075 3'.
								W90388	zh75f08.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417927 3'
								F03786	H. sapiens partial cDNA sequence; clone c-29h08.
72	CATGTACTGTACTTC	H817952	12	2	0	0	0	U14631	Human 11 beta-hydroxysteroid dehydrogenase type II
73	CATGCCCTTGCACTC	H360008	11	6	0	3	3	T41121	ya31a06.s5 Homo sapiens cDNA clone 62194 3' contains Alu repetitive element,.
74	CATGCGGTGGGACCA	H440966	11	4	0	2	0		Unknown
75	CATGGCCCCCAACCA	H611590	11	2	0	0	0		Unknown
76	CATGGCCGGCGCTC	H616862	11	2	0	0	0	Z58486	Unknown
77	CATGGGAGGCGCTCA	H666014	11	1	0	0	0		Unknown

78	CATGTCCCCGTTACA	H874226	11	11	0	0	0	W68073	zd42c12.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 343318 3' similar to contains Alu repetitive element;

Table 4 - Transcripts increased in pancreas_cancer -

SAGE Tags elevated only in Pancreatic Tumor

NC Normal Colon
Tu Colon Tumor
CC Colon Cancer Cell Line
PT Pancreatic Tumor
PC Pancreatic Cell Line

	Tag Sequence	Tag Number	NC	Tu	CC	PT	PC		Accession	Gene Name
1	CATGAAAGCAAACCA	H9222	0	6	1	3	11	Examples	R38305	yh95b04.s1 Homo sapiens cDNA clone 137455 3'
									AA126719	zk95b03.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 490541 3'
									AA044296	zk51c03.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486340 3'
									AA131586	zl33c08.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 503726 3'
2	CATGAAAGCAGTTTA	H9408	1	5	2	21	3	Examples	AA157983	zo71h12.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592391 3'
									AA292929	zt54e04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 726174 3'
									AA159306	zo78c07.s1 Stratagene pancreas (#937208) Homo zo78c07.s1 Stratagene pancreas (#937208) Homo
									R54012	yj70h01.s1 Homo sapiens cDNA clone 154129 3'
									T62936	yb99f08.s1 Homo sapiens cDNA clone 79335 3'
3	CATGAAAGCGGGGCT	H9898	0	0	0	0	13	Examples	X52426	H. sapiens mRNA for cytokeratin 13
4	CATGAAATCCTGGGT	H13803	0	1	1	16	2	Examples	X51698	H.sapiens spasmolytic polypeptide (SP) mRNA.
5	CATGAAATGGACAAC	H14865	0	0	1	0	13	Examples	N70419	za61d12.s1 Homo sapiens cDNA clone 297047 3'
									AA411599	zv16g01.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 753840 5'
									AA410508	zv16g01.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 753840 3'
6	CATGAACCAGTTTGT	H21247	1	1	6	8	13	Examples	AA115723	zl86g12.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511558 3'
									AA132875	zo19e04.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 587358 3'
									AA147677	zo44a06.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 589714 3'

									AA206883	zq81h12.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 648071 3'
								Examples	R51318	yg72f03.s1 Homo sapiens cDNA clone 38681 3'
7	CATGAACTCTTGAAG	H30689	3	7	13	13	17		T35270	EST82235 Homo sapiens cDNA 3' end similar to None
									AA412071	zt65h12.s1 Soares testis NHT Homo sapiens cDNA clone 727271 3'
								Examples	N63154	yz37f12.s1 Homo sapiens cDNA clone 285263 3'
8	CATGAACTGCTTCAA	H31221	7	6	8	6	130		T87236	yc81h04.s1 Homo sapiens cDNA clone 22603 3'
									AA150720	zl46f04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 5049
									AA045773	zl68b12.s1 Stratagene colon (#937204) Homo sapiens
								Examples	X07819	Human pump-1 mRNA homolog. to metalloproteinase,
9	CATGAACTTGGCCAT	H32405	0	0	0	8	11		L22523	Human matrilysin gene, exon 5
10	CATGAAGATCCCCGC	H36183	5	10	14	12	23	Examples	R72650	yj95e05.s1 Homo sapiens cDNA clone 156512 3'
									W70287	zd58e02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344858 3' similar to SW:CUTA_ECOLI P36654 PERIPLASMIC DIVALENT CATION TOLERANCE PROTEIN CUTA
									R72650	yj95e05.s1 Homo sapiens cDNA clone 156512 3' similar to SP:CYCY_ECOLI P36654 C-TYPE CYTOCHROME BIOGENESIS PROTEIN CYCY
									AA181976	zp61a11.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 624668 3' similar to SW:CUTA_ECOLI P36654 PERIPLASMIC DIVALENT CATION TOLERANCE PROTEIN CUTA
11	CATGAAGGGAGGGTC	H43180	6	3	8	15	41	Examples	U46751	Human phosphotyrosine independent ligand p62 for tthe Lck SH2 domain mRNA, complete cds
12	CATGAAGTTGCTATT	H48756	10	9	18	31	27	Examples	J03077	Human co-beta glucosidase (proactivator) mRNA
									M86181	Human prosaposin (PSAP) gene
									D00422	Human sphingolipid activator proteins, mRNA
									J03015	Homo sapiens sphingolipid activator protein 1 mRNA
									M60255	Human mutant cerebroside sulfate activator protein
13	CATGAATGAAAAAAA	H57345	0	1	5	2	10	No Match		
14	CATGACAAACTGTGG	H66031	17	4	24	5	60	Examples	N22375	yw37d01.s1 Homo sapiens cDNA clone 254401 3'
									AA084643	zn20e01.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 547992 3'

									AA279290	zs84a06.s1 Soares NbHTGBC Homo sapiens cDNA clone 704146 3'
									AA046253	zf12a02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 376682 3'
15	CATGACAACTCAATA	H67396	2	7	7	16	37	Examples	Z58016	H.sapiens CpG DNA, clone 26c7,
									AA151668	zo29c02.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 588290 3' similar to SW:BI3_MOUSE P28662 BRAIN PROTEIN I3
									W02958	za07e06.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone 291874 5'
16	CATGACACCCTGTGC	H71151	0	1	0	2	14	Examples	AA1556464	zo70e05.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592256 3'
									AA025673	ze90h09.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366305 3'
									N70895	za89h12.s1 Homo sapiens cDNA clone 299783 3'
17	CATGACCATTGGATT	H85924	0	8	5	13	4	Examples	X02491	Human interferon-inducible mRNA (cDNA 9-27): membrane
									J04164	Human interferon-inducible protein 9-27 mRNA
									X84958	H.sapiens mRNA for interferon-induced 17kDa membra
18	CATGACCCTTTAACA	H90050	1	4	2	13	7	Examples	X56841	H.sapiens HLA-E gene.
									X64879	H.sapiens mRNA for HLA-E heavy chain (exons 4 - 7)
19	CATGACCGCCGTGGT	H91579	49	22	45	70	94	Examples	M21186	Human neutrophil cytochrome b light chain p22A
									M61107	Human p22-phox (CYBA) gene, exons 3 and 4
20	CATGACCTGTGACCA	H97158	0	3	0	28	17	Examples	D00244	Human Pro-urokinase gene,
									K02286	Human urokinase gene, 3' end
									M15476	Human pro-urokinase mRNA, complete cds
									X02419	Human uPA gene for urokinase-plasminogen activator
21	CATGACGCCCTGCTC	H103912	0	1	0	11	2	Examples	L08835	Human myotonic dystrophy kinase (DM kinase) gene
									M87313	Homo sapiens myotonin protein kinase (DM) mRNA
22	CATGACGTGGTGATG	H113380	2	4	4	5	20	Examples	H44451	yo75f06.s1 Homo sapiens cDNA clone 183779 3'
									AA157329	zo42f07.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 589573 3' similar to SW:L10K_RAT Q05310 LEYDIG CELL TUMOR 10 KD PROTEIN
									W46455	zc32g06.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 324058 3' similar to SW:L10K_RAT Q05310 LEYDIG CELL TUMOR 10 KD PROTEIN

23	CATGACTCAGCCCGG	H119383	0	0	3	21	3	Examples	M92357	Homo sapiens B94 protein mRNA, complete cds.
24	CATGACTGAGGAAAG	H123521	0	0	0	53	22	Examples	X64875	H.sapiens mRNA for insulin-like growth factor binding protein 3
									M31159	Human growth hormone-dependent insulin-like growth factor binding protein 3
									M35878	Human insulin-like growth factor-binding protein-3
									S56205	insulin-like growth factor binding protein 3 {3' region}
25	CATGACTGCCCGCTG	H124264	1	0	0	22	9	Examples	U65932	Human extracellular matrix protein 1 (ECM1) mRNA
									U65937	Human extracellular matrix protein 1 (ECM1) gene, exon 9
26	CATGACTGTATTTTC	H126208	3	4	9	2	22	Examples	AA148916	zo03f09.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 566633 3'
									AA129137	zo12a11.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 586652 3'
									AA115437	zl85g09.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511456 3'
									AA126967	zl87e07.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511620 3'
27	CATGAGCACTGCAGC	H149395	1	2	6	3	16	Examples	R24613	yh36c03.r1 Homo sapiens cDNA clone 131812
28	CATGAGCAGGAGCGT	H150055	1	0	0	0	15	Examples	H43243	yp05e05.r1 Homo sapiens cDNA clone 186560 5'
29	CATGAGCTGTATTCT	H162622	0	2	0	1	11	Examples	X54942	H.sapiens ckshs2 mRNA for Cks1 protein homologue
30	CATGAGGATGACCCC	H167446	1	7	12	10	13	Examples	AA044081	zk50g07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486300 3'
									AA044211	zk50g07.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 486300 5' similar to PIR:A40533 A40533 cAMP-dependent protein kinase major membrane substrate
31	CATGAGGTCTTCAAT	H178129	4	2	0	60	2	Examples	X14787	Class A, Human mRNA for thrombospondin.
32	CATGAGGTGCGGGGG	H178603	0	2	2	1	11	Examples	R27738	yh64f11.s1 Homo sapiens cDNA clone 134541 3'
									H00276	yj22f12.s1 Homo sapiens cDNA clone 149519 3' similar to SP:ZK637.5 CE00436 ARSA
33	CATGAGTATCTGGGA	H183787	3	3	1	15	73	Examples	AA076235	zm19d07.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526093 3'
									H13159	yj16c04.s1 Homo sapiens cDNA clone 148902 3'
									AA146632	zo71e11.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592364 3'
34	CATGATACTTTAATT	H204740	1	0	3	18	9	Examples	X80062	H.sapiens SA mRNA.
									U01691	Human annexin V (ANX5) gene

									X12454	Human mRNA for vascular anticoagulant
									M18366	Human placental anticoagulant protein (PAP) mRNA
									M21731	Human lipocortin-V mRNA, complete cds
									J03745	Human endonexin II mRNA, complete cds
35	CATGATCAAGAATCC	H213518	2	1	5	25	1	Examples	J03909	GAMMA-INTERFERON-INDUCIBLE PROTEIN IP-30 PRECURSOR (HUMAN)
									aa383911	EST97384 Thymus II Homo sapiens cDNA 3' end similar to interferon, gamma transducer 1
36	CATGATCAAGGGTGT	H213679	12	9	25	12	156	Examples	U09953	Human ribosomal protein L9 mRNA
									U21138	Human ribosomal protein L9 mRNA, complete cds
									D14531	Human mRNA for human homologue of rat ribosomal protein
37	CATGATCAAGTTCGA	H213751	0	2	8	3	10	Examples	AA063259	zm03a05.s1 Stratagene corneal stroma (#937222) Homo sapiens cDNA clone 513008 3'
38	CATGATCCGGCGCCA	H219750	16	7	14	12	40	Examples	L42856	RNA polymerase II transcription factor SIII p18 subunit mRNA
39	CATGATGAAACTTCG	H229502	1	0	0	17	4	Examples	Z59242	H.sapiens CpG DNA, clone 13a10, reverse read cpg1
40	CATGATGCGAAAGGC	H235531	2	3	12	3	22	Examples	Z25820	H.sapiens mRNA for mitochondrial dodecenoyl-CoA dehydrogenase
									L24774	Homo sapiens delta3, delta2-CoA-isomerase mRNA
41	CATGATGTCTTCGTT	H243676	0	0	1	0	14	Examples	M84711	40S RIBOSOMAL PROTEIN S3A (HUMAN)
42	CATGATGTCTTTTCT	H243710	1	2	1	14	2	Examples	M62403	Human insulin-like growth factor binding protein 4
									U20982	Human insulin-like growth factor binding protein-4 (IGFBP4) gene, promoter and complete cds
43	CATGATGTGTAACGA	H244487	0	4	5	44	94	Examples	Z33457	H.sapiens mts1 gene.
									M80563	Human CAPL protein mRNA, complete cds
44	CATGCAACTTAAAGC	H270083	0	1	2	10	1	Examples	N23207	yx70b09.s1 Homo sapiens cDNA clone 267065 3' similar to gb:L12350 THROMBOSPONDIN 2 PRECURSOR (HUMAN)
45	CATGCACCTGTCCTT	H286424	0	4	2	10	1	Examples	AA285023	zt25e11.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 714188 3' similar to gb:M33680 CD81 ANTIGEN (HUMAN)
									M33680	CD81 antigen
46	CATGCACTCAATAAA	H291889	0	0	2	3	19	Examples	D78203	Neurosin
									U62801	protease M

47	CATGCAGCCTGGGGC	H300971	0	0	0	0	10	Examples	AA149942	zo68d04.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592039 3' similar to TR:E218488 E218488 TRYPTASE
48	CATGCAGCGCGCCCT	H301462	4	11	12	10	21	Examples	AA187553	zp66b09.r1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 625145 5' similar to gb:M16937 HOMEOBOX PROTEIN HOX-B7 (HUMAN);contains element MER22 repetitive element
									M16937	Homeobox protein HOX-B7
49	CATGCAGGTTGTCCT	H307126	0	0	4	0	10	No Match		
50	CATGCAGTCTCTCAA	H309109	2	6	6	2	17	Examples	U14972	Human ribosomal protein S10 mRNA
51	CATGCATCCCGTGAC	H316857	0	3	3	3	13	Examples	U27293	Human leukotriene A4 hydrolase gene
									J03459	Human leukotriene A-4 hydrolase mRNA, complete cds
									J02959	Human leukotriene A-4 hydrolase mRNA, complete cds
52	CATGCATTCCTCCTT	H325080	0	2	5	13	3	Examples	X82434	H.sapiens mRNA for emerin
53	CATGCCACCCCCACC	H333138	3	7	17	18	2	Examples	M88338	Human serum constituent protein (MSE55) mRNA
54	CATGCCAGTGGCCCG	H339606	23	11	37	22	56	Examples	U14971	Human ribosomal protein S9 mRNA
55	CATGCCATTTTCTGG	H344031	0	2	6	1	10	Examples	L01697	Homo sapiens alpha-1 type XV collagen mRNA
56	CATGCCCAAGCTAGC	H344691	19	8	8	18	44	Examples	X54079	Human mRNA for heat shock protein HSP27.
									Z23090	H.sapiens mRNA for 28 kDa heat shock protein
									X16477	Human mRNA fragment for estrogen-regulated 24k protein
									S74571	estrogen receptor-related protein=27-kda heat shock protein
57	CATGCCCATCCGAAA	H347489	20	15	43	19	61	Examples	X69392	H.sapiens mRNA for ribosomal protein L26.
									L07287	Human ribosomal protein L26 (RPL26) gene
58	CATGCCCCCTGCAGA	H350099	0	1	6	14	25	Examples	U40434	Human mesothelin or CAK1 antigen precursor mRNA
									D49441	Human mRNA for pre-pro-megakaryocyte potentiating factor, complete cds.
59	CATGCCCGCATAGAT	H353481	0	0	0	8	11	Examples	U12819	Human p16-INK4 (p16) gene
									U38945	Human hypothetical 18.1 kDa protein (CDKN2A) mRNA
									S69804	MTS1=multiple tumor suppressor 1/cyclin-dependent kinase 4 inhibitor p16
									S69822	CDK4I=cyclin-dependent kinase 4 inhibitor
									S78535	tumor suppressor gene, P16/MTS1/CDKN2=cell cycle cycle negative regulator beta form
60	CATGCCCTCCTGGGG	H357867	8	2	5	14	34	Examples	Z47319	H.sapiens mRNA for expressed sequence tag (clone 21fi7I19)

									AA398406	zt60h12.s1 Soares testis NHT Homo sapiens cDNA clone 726791 3'
61	CATGCCGGCCCTACC	H370034	4	4	1	14	19	Examples	U21049	Human DD96 mRNA
62	CATGCCTGGTCCCAA	H387925	0	2	1	30	99	Examples	X03212	KERATIN, TYPE II CYTOSKELETAL 7
									AA187637	zp73f01.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 625849 3'
63	CATGCCTTTGAACAG	H392709	5	3	6	2	23	Examples	AA176457	zp35g11.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 611492 3' similar to TR:G663269 G663269 BOLA
									AA176541	zp35e11.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 611468 3' similar to TR:G663269 G663269 BOLA.
64	CATGCGCCGACGATG	H415844	21	13	45	75	7	Examples	X02492	Human interferon-inducible mRNA fragment
65	CATGCTCAACAGCAA	H475429	2	5	10	6	17	Examples	T53402	ya88g05.s1 Homo sapiens cDNA clone 68792 3'
									W69493	zd47g08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 343838 3' similar to PIR:S24168 S24168 hypothetical protein - human
66	CATGCTCAACCCCCC	H475478	1	4	2	23	1	Examples	X13916	Human mRNA for LDL-receptor related protein
67	CATGCTGAGAAACTG	H493576	2	3	1	8	18	Examples	X80335	H.sapiens (24) Ferritin H pseudogene.
68	CATGCTGAGTCTCCC	H494454	1	4	4	21	13	Examples	X04828	Human mRNA for G(i) protein alpha-subunit
69	CATGCTGCTATACGA	H498887	16	30	28	30	44	Examples	U14966	Human ribosomal protein L5 mRNA
70	CATGCTGCTGAGTGA	H499247	1	3	4	13	13	Examples	T90665	yd41g08.s1 Homo sapiens cDNA clone 110846 3'
									AA338799	EST43791 Fetal brain I Homo sapiens cDNA 3' end similar to steroid hormone receptor hERR1
									H97236	yv98b06.s1 Homo sapiens cDNA clone 250739 3'
71	CATGCTGGCGCCGAT	H501337	0	0	4	0	10	Examples	C14084	Human fetal brain cDNA 3'-end GEN-018D10
72	CATGCTTCCAGCTAA	H513181	64	23	36	53	104	Examples	D00017	Human lipocortin II mRNA
73	CATGCTTCCTTGCCT	H514022	0	3	4	89	7	Examples	Z19574	H.sapiens gene for cytokeratin 17.
									X62571	H.sapiens mRNA for keratin-related protein
									X05803	Human radiated keratinocyte mRNA 266
74	CATGCTTTCTTCCCT	H522198	0	2	1	16	4	Examples	X79067	H.sapiens ERF-1 mRNA 3' end.
75	CATGGAAAAAAAAAA	H524289	7	14	21	26	37	Examples	X51779	Human mRNA containing an Alu repeat
									X82240	H.sapiens mRNA for Tcell leukemia/lymphoma 1
76	CATGGAAACAAGATG	H525348	4	7	14	8	22	Examples	V00572	Human mRNA encoding phosphoglycerate kinase.
									D29018	Human keratinocyte cDNA, clone 001
									L00160	Human phosphoglycerate kinase (pgk) mRNA
77	CATGGAAATACAGTT	H527436	49	35	10	100	36	Examples	X05344	Human mRNA for cathepsin D

78	CATGGAAATGATGAG	H527929	4	7	5	14	26	Examples	M11233	Human cathepsin D mRNA, complete cds
									T90296	yd42f03.s1 Homo sapiens cDNA clone 110909 3' similar to SP:R151.9 CE00827
									AA320942	EST23523 Adipose tissue, brown Homo sapiens cDNA 3' end
79	CATGGAAGATGTGTG	H533436	3	7	16	6	28	Examples	AA181811	zp64f07.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 624997 3'
									AA148508	zl06c06.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491530 3' similar to WP:ZK652.2 CE00448
80	CATGGAATTTTATAA	H540621	6	3	10	9	28	Examples	L21950	Human peripheral benzodiazepine receptor related mRNA
									M36035	Human peripheral benzodiazepine receptor (hpbs) mRNA
81	CATGGACAAAAAAAA	H540673	1	2	10	3	17	No Match		
82	CATGGACCACCTTTA	H545152	0	1	0	11	2	Examples	U19718	Human microfibril-associated glycoprotein (MFAP2).
83	CATGGACCAGGCCCT	H545430	0	3	0	20	18	Examples	M75165	H.sapiens epithelial tropomyosin (TM1) mRNA
									M12125	Human fibroblast muscle-type tropomyosin mRNA
									M74817	Human tropomyosin-1 (TM-beta) mRNA, complete cds
84	CATGGACCCCAAGGC	H546059	2	5	9	16	10	Examples	M74092	Human cyclin mRNA
85	CATGGACCCTGCCCT	H546710	31	36	20	71	65	Examples	L37033	Homo sapiens FK-506 binding protein homologue
86	CATGGACCTATCTCT	H548062	0	1	0	13	1	Examples	N90046	zb37g02.s1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 305810 3'
									AA115048	zl06a10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491514 3'
87	CATGGACGGCGCAGG	H551315	3	4	5	32	3	Examples	M63193	Human platelet-derived endothelial cell growth factor
88	CATGGACTCTCTGTT	H554876	1	4	3	0	14	Examples	M61764	Human gamma-tubulin mRNA,
89	CATGGAGAGCTTTGC	H559615	0	0	0	2	10	Examples	D17793	Human mRNA (HA1753) for ORF
90	CATGGAGAGTGTCTG	H560056	0	5	8	32	11	Examples	S68252	TIMP-1=metalloproteinase inhibitor
									X02598	EPA glycoprotein (erythroid-potentiating activity)
									X03124	tissue inhibitor of metalloproteinase 2
91	CATGGAGCAGGATGA	H561807	0	0	0	1	12	No Match		
92	CATGGAGGGAGTTCC	H567486	1	1	0	4	13	Examples	AA214523	zr89c01.s1 Soares NbHTGBC Homo sapiens cDNA clone 682848 3'
									N30324	yw75d01.s1 Homo sapiens cDNA clone 258049 3'
93	CATGGAGTCCGGAGC	H570787	0	0	2	1	10	Examples	X70070	H.sapiens mRNA for neurotensin receptor.
94	CATGGAGTTATGTTG	H572656	0	0	3	0	10	Examples	H57673	yr27a10.s1 Homo sapiens cDNA clone 206490 3'

									W94333	ze12c08.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 358766 3' similar to SW:YA94_SCHPO Q09783 HYPOTHETICAL 11.4 KD PROTEIN C13G6.04 IN CHROMOSOME I
95	CATGGAGTTCGACCT	H572806	7	3	7	15	29	No Match		
96	CATGGATTAAGTGAG	H585913	3	5	2	2	19	Examples	AA046631	zk72d06.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 488363 3'
									R91942	yq06g03.s1 Homo sapiens cDNA clone 196180 3'
									AA040439	zk46c12.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 485878 3'
97	CATGGATTGAACCTC	H587800	1	0	5	1	12	Examples	U60205	methyl sterol oxidase (ERG25)
98	CATGGCAAAAAAAAA	H589825	17	13	29	73	38	No Match		
99	CATGGCATTTAAATA	H605956	2	10	8	3	55	Examples	X60489	Human mRNA for elongation factor-1-beta.
									X60656	H.sapiens mRNA for elongation factor 1-beta
100	CATGGCCAACAACGA	H606471	0	0	0	12	1	Examples	U08021	Human nicotinamide N-methyltransferase (NNMT) mRNA, 0
101	CATGGCCCCCAATAA	H611597	1	4	1	47	9	Examples	X15256	Human mRNA for 14kDa beta-galactoside-binding lectin
									X14829	Human mRNA for beta-galactoside-binding lectin
									J04456	Human 14 kd lectin mRNA, complete cds
									S44881	HL14=beta-galactoside binding protein
102	CATGGCCGCTACTTC	H616224	0	0	1	3	16	Examples	AA054483	zk82d04.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 489319 5' similar to contains Alu repetitive element
103	CATGGCCGTCGGAGG	H617891	8	5	2	44	3	Examples	AA243725	zr68g12.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 668614 3' similar to gb:X02492 INTERFERON-INDUCED PROTEIN 6-16 PRECURSOR (HUMAN)
104	CATGGCCTACCCGAG	H618841	0	4	4	23	39	Examples	X13425	Human mRNA for pancreatic carcinoma marker GA733-1, 0
105	CATGGCGGGGTGGAG	H633577	3	8	5	27	6	Examples	AA136985	zl02b03.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491117 3'
106	CATGGCTCAGCTGGA	H643707	12	29	24	35	35	Examples	AA053346	zl70h04.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 510007 3' similar to gb:Z21507 ELONGATION FACTOR 1-DELTA
107	CATGGCTTTTCAGAC	H655177	1	6	7	13	10	Examples	U43368	Human VEGF related factor isoform VRF186 precursor, 0
									U52819	Human vascular endothelial growth factor B 186
108	CATGGGAAAAAAAAA	H655361	11	8	30	16	38	Examples	M38259	Human cytochrome c oxidase subunit VIb
									M60748	Human histone H1 (H1F4) gene, complete cds

									M73239	Human (clone SF1) hepatocyte growth factor (HGF)
									M73240	Human (clone SF2) hepatocyte growth factor (HGF)
109	CATGGGAAAAGTGGT	H655547	18	13	3	70	1	Examples	X02920	Human mRNA for alpha 1-antitrypsin carboxyterminal, 0
									X01683	Human mRNA for alpha 1-antitrypsin
									V00496	Human messenger RNA for alpha-1-antitrypsin
									J00067	Human alpha-1 antitrypsin gene, 3' end
110	CATGGGAAGGGAGGC	H658059	0	0	4	6	16	Examples	AA127040	zl22b01.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 502633 3'
									W81387	zd86f06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347555 3'
									H45477	yo72h08.s1 Homo sapiens cDNA clone 183519 3'
111	CATGGGAGTCATTGT	H666943	6	5	6	10	32	Examples	D26598	Human mRNA for proteasome subunit HsC10-II. , 0
112	CATGGGAGTGTGCGT	H667367	0	0	1	1	10	Examples	N74310	za78c01.s1 Homo sapiens cDNA clone 298656 3'
									H92750	yt92e01.s1 Homo sapiens cDNA clone 231768 3'
									T24084	seq2272 Homo sapiens cDNA clone ssb4HB3MA(extended-ft-6) 3'
113	CATGGGATTGTCTGG	H671455	3	7	13	5	21	Examples	X17567	H.sapiens RNA for snRNP protein B
									M34081	Human small nuclear ribonucleoprotein particle SmB
114	CATGGGCCCCTCACC	H677330	0	0	2	9	22	Examples	M69054	Human insulin-like growth factor binding protein 6, 0
									M62402	Human insulin-like growth factor binding protein 6
115	CATGGGCCCTCTGAG	H677753	0	1	4	7	14	Examples	N74323	za78d08.s1 Homo sapiens cDNA clone 298671 3'
									H46766	yo18f08.s1 Homo sapiens cDNA clone 178311 3'
									H41102	yn88a08.s1 Homo sapiens cDNA clone 175478 3'
116	CATGGGCTGGTCTGG	H686815	0	1	3	13	22	Examples	AA074777	zm84b09.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 544601 3'
									AA062735	zm04a04.s1 Stratagene corneal stroma (#937222) Homo sapiens cDNA clone 513102 3'
									AA112905	zm63f12.s1 Stratagene fibroblast (#937212) Homo sapiens cDNA clone 530351 3'
117	CATGGGGAAGCAGAT	H688713	25	7	9	0	72	No Match		
118	CATGGGGAGGGGTGG	H690863	2	3	1	16	2	No Match		
119	CATGGGGAGGTAGCA	H690890	1	0	1	14	1	No Match		
120	CATGGGGCATCTCTT	H693112	1	1	3	39	2	Examples	V00523	Human mRNA for histocompatibility antigen HLA-DR
									X00274	Human gene for HLA-DR alpha heavy chain a class II
									K01171	Human HLA-DR alpha-chain mRNA

									J00202	human hla-dr heavy chain gene; 3' flank
121	CATGGGTGGGGAGAT	H715401	1	4	10	10	14	Examples	U18009	Human chromosome 17q21 mRNA clone LF113.
									T33413	EST57778 Homo sapiens cDNA 3' end similar to None
									T33339	EST57474 Homo sapiens cDNA 3' end similar to None
122	CATGGTACTGTAGCA	H728778	3	3	1	16	30	Examples	M59911	Human integrin alpha-3 chain mRNA
123	CATGGTACTGTGGCT	H728810	23	10	16	15	50	Examples	X87689	H.sapiens mRNA for putative p64 CLCP protein
124	CATGGTCAAAATTTC	H737344	0	0	0	10	1	Examples	L12350	Human thrombospondin 2 (THBS2) mRNA
125	CATGGTCTGGGGCTT	H752296	25	35	45	76	29	Examples	D21261	Human mRNA (HA1756) for ORF
									D29543	Human keratinocyte cDNA, clone 686
126	CATGGTCTGTGAGAG	H752521	0	5	7	12	2	Examples	H51290	yp07a05.s1 Homo sapiens cDNA clone 186704 3'
									N20338	yx44g12.s1 Homo sapiens cDNA clone 264646 3'
									AA158271	zo76e09.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592840 3'
127	CATGGTCTGTGCAGG	H752531	0	0	0	1	13	No Match		
128	CATGGTCTTGAAGCC	H753162	0	1	2	1	10	No Match		
129	CATGGTGAAGGCAGT	H754323	25	14	42	15	89	Examples	X87373	Class C, H.sapiens RPS3a gene
130	CATGGTGAATGACGG	H754567	0	2	8	1	10	Examples	X08058	GLUTATHIONE S-TRANSFERASE P (HUMAN)
131	CATGGTGCGGAGGAC	H760361	0	3	2	11	25	Examples	X51439	Human mRNA for serum amyloid A (SAA) protein
132	CATGGTGCTGGAGAA	H761481	2	9	9	13	26	Examples	U15008	Human SnRNP core protein Sm D2 mRNA
133	CATGGTGGAGGGCAC	H762533	1	1	3	6	34	Examples	U62800	Cystatin M (CST6)
134	CATGGTGGTACAGGA	H765003	14	17	15	39	30	Examples	H46430	yo12h12.s1 Homo sapiens cDNA clone 177767 3'
									AA047563	zf13a06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 376786 3'
									AA130701	zo13f02.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 586779 3'
135	CATGGTTCACTGCAG	H774629	0	2	1	13	3	Examples	X59288	H.sapiens gene for intercellular adhesion molecule
									M24283	Human major group rhinovirus receptor (HRV) mRNA
									J03132	Human intercellular adhesion molecule-1 (ICAM-1)
									M55100	Human cell surface glycoprotein P3.58 mRNA
136	CATGGTTGTCTTTGG	H781823	1	1	6	30	24	Examples	K02765	Human complement component C3 mRNA, alpha and beta
137	CATGGTTGTGGTTAA	H782013	178	110	14	340	139	Examples	M17987	Human beta-2-microglobulin gene
138	CATGGTTTAAATCGA	H782391	1	6	12	4	14	Examples	D00760	Human mRNA for proteasome subunit HC3
139	CATGTAAGGCTTAAC	H797169	0	0	6	1	12	Examples	X57025	INSULIN-LIKE GROWTH FACTOR IA PRECURSOR (HUMAN)
140	CATGTAATTTTGGAA	H802793	0	2	5	2	10	No Match		

	CATGTAATTTTGGAT	H802793						No Match		
141	CATGTACATTTTCAT	H806901	1	4	2	3	14	Examples	X85373	H.sapiens mRNA for Sm protein G
142	CATGTACCCCGTACA	H808370	0	1	4	0	10	No Match		
143	CATGTACCCTTCTAT	H808925	0	0	0	17	7	No Match		
144	CATGTAGGAAAGTAA	H827437	1	0	5	5	24	Examples	J02931	Human placental tissue factor (two forms) mRNA
									M16553	Human tissue factor mRNA, complete cds
									M27436	Human tissue factor gene, complete cds
145	CATGTAGGTTGTCTA	H831416	49	61	61	89	130	Examples	X64899	H.sapiens mRNA homologous to mouse P21 mRNA.
									X16064	Human mRNA for translationally controlled tumor protein
									L13806	Homo sapiens (clone 04) translationally controlled tumor protein
146	CATGTATATTTTCTC	H839672	1	0	3	8	16	Examples	M98479	Human transglutaminase mRNA
147	CATGTATTTTCTGCC	H851834	0	1	2	16	3	Examples	D12149	Human HepG2 3'-directed MboI cDNA, clone s247
148	CATGTCACAAGCAAA	H856209	10	28	27	24	48	Examples	X80909	H.sapiens alpha NAC mRNA
149	CATGTCCAAATCGAT	H868569	0	1	0	43	17	Examples	X56134	Human mRNA for vimentin.
									Z19554	H.sapiens vimentin gene
									M14144	Human vimentin gene, complete cds
									M25246	Human vimentin (HuVim3) mRNA, 3' end
150	CATGTCCACTGGCCT	H870310	0	0	1	12	2	Examples	N92906	zb57a08.s1 Homo sapiens cDNA clone 307670 3'
									T17488	NIB978 Normalized infant brain, Bento Soares Homo sapiens cDNA 3'end
									AA349906	EST56900 Infant brain Homo sapiens cDNA 3' end
151	CATGTCCATCTGTTG	H871920	6	6	10	25	5	Examples	X67016	H.sapiens mRNA for amphiglycan
									D13292	Human mRNA for ryudocan core protein
152	CATGTCGTCTTTATC	H899060	2	5	15	1	69	Examples	M77233	Human ribosomal protein S7 mRNA
153	CATGTCTCTGATGCT	H908858	1	5	2	46	19	Examples	S48568	tissue inhibitor of metalloproteinase 2 (3'-end region)
154	CATGTCTTGTAACTG	H916232	0	4	3	1	13	Examples	N71680	yz93b03.s1 Homo sapiens cDNA clone 290573 3'
155	CATGTCTTGTGCATA	H916372	14	22	15	20	45	Examples	X03083	Human lactate dehydrogenase-A gene
									X02152	Human mRNA for lactate dehydrogenase-A
									X02153	Human pseudogene for lactate dehydrogenase-A
156	CATGTGAAGTCACTG	H920392	1	1	6	0	16	No Match		
157	CATGTGAAGTTATAC	H920525	0	1	3	6	11	Examples	X07979	CTGTGG, Class A, Human mRNA for fibronectin receptor beta subunit.

158	CATGTGATGTCTGGT	H932731	0	8	3	11	12	Examples	AA027860	zk05h07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 469693 3'
159	CATGTGCCATCTGTA	H938876	1	3	7	3	16	Examples	M25753	G2/MITOTIC-SPECIFIC CYCLIN B1 (HUMAN)
									T60151	yc22c04.s1 Homo sapiens cDNA clone 81414 3'
									R67969	yi29g08.s1 Homo sapiens cDNA clone 140702 3'
			.							
160	CATGTGCCCTCAAAA	H939841	11	13	3	13	43	Examples	AA169614	zo91f03.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 594269 3' similar to SW:NGAL_HUMAN P80188 NEUTROPHIL GELATINASE-ASSOCIATED LIPOCALIN PRECURSOR
161	CATGTGCCCTCAGAA	H939849	3	4	0	11	19	Examples	N79823	zb15d08.s1 Homo sapiens cDNA clone 302127 3' similar to SW:NGAL_HUMAN P80188 NEUTROPHIL GELATINASE-ASSOCIATED LIPOCALIN PRECURSOR
162	CATGTGCCCTCAGGA	H939851	13	31	10	25	83	Examples	AA075896	zm90h04.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 545239 3' similar to SW:NGAL_HUMAN P80188 NEUTROPHIL GELATINASE-ASSOCIATED LIPOCALIN PRECURSOR
162	CATGTGCCCTCAGGC	H920392						No Match		
163	CATGTGCCTTACTTT	H941856	0	3	1	2	12	Examples	AA100279	zl81c07.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511044 3'
164	CATGTGCGCTGGCCC	H944038	2	5	2	17	3	No Match		
165	CATGTGCTTCATCTG	H949560	2	6	6	4	16	Examples	AA029262	zk10a01.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470088 3'
									N54281	yv66e10.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone 247722 3'
									AA114075	zn76c02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 564098 3'
166	CATGTGGAGTGGAGG	H953251	18	15	7	22	48	Examples	L76200	Homo sapiens guanylate kinase (GUK1) mRNA
167	CATGTGGCCCCAGGT	H955723	0	3	3	37	4	Examples	X00570	Human mRNA for precursor of apolipoprotein CI
168	CATGTGGGTGAGCCA	H962086	13	15	13	76	27	Examples	L16510	Homo sapiens cathepsin B mRNA
									M14221	Human cathepsin B proteinase mRNA, complete cds
169	CATGTGTGAGCCCCT	H975446	3	3	3	22	8	Examples	L35240	Human enigma gene
170	CATGTGTGCTAAATG	H976644	8	21	26	18	50	Examples	L38941	Homo sapiens ribosomal protein L34 (RPL34) mRNA
171	CATGTGTGTGTTTGT	H978687	6	7	16	25	15	Examples	X03473	Human gene for histone H1(0).
172	CATGTTATGGATCTC	H997944	0	1	1	21	1	Examples	AA034505	zk23g08.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 471422 3'

									AA235464	zt31b06.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 723923 3'
									AA037024	zk30c10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 472050 3'
173	CATGTTCATTGTAGA	H1003443	0	7	0	10	3	Examples	H53629	yu38d04.s1 Homo sapiens cDNA clone 236071 3'
									T06706	EST04595 Homo sapiens cDNA clone HFBDX32
									T16635	NIB1599 Normalized infant brain, Bento Soares Homo sapiens cDNA 3'end similar to EST04595 H. sapiens cDNA clone HFBDX32
174	CATGTTCTGTGAATC	H1014660	3	4	3	24	5	Examples	AA026678	ze97h02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366963 3'
									AA280283	zt05a03.s1 Soares NbHTGBC Homo sapiens cDNA clone 712204 3'
									H10141	ym05a09.s1 Homo sapiens cDNA clone 46675 3'
175	CATGTTGCCCCCGTG	H1021276	0	0	0	8	17	Examples	X66029	H.sapiens mRNA for tyrosine kinase receptor.
176	CATGTTGCTGACTTT	H1023520	1	5	1	33	1	Examples	X15880	Human mRNA for collagen VI alpha-1
									X72414	H.sapiens gene for glutaminyl-tRNA synthetase
177	CATGTTGGAGATCTC	H1024568	4	11	16	10	24	Examples	AA044568	zk73h10.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 488515 3'
									N71899	yz36b07.s1 Homo sapiens cDNA clone 285109 3'
									AA400793	zt71g03.s1 Soares testis NHT Homo sapiens cDNA clone 727828 3'
178	CATGTTGGGGTTTCC	H1026814	202	75	84	235	369	Examples	X80336	H.sapiens (5) Ferritin H pseudogene.
									X00318	Human mRNA for apoferritin H chain type
									X03488	Human apoferritin H gene exons 2-4
									M97164	Human ferritin heavy chain mRNA, complete cds
									L20941	Human ferritin heavy chain mRNA, complete cds
179	CATGTTGGTGAAGGA	H1027595	98	106	17	183	107	Examples	X02493	Human interferon-inducible mRNA (cDNA 6-26).
									M11948	Human promyelocytic leukemia cell mRNA
									M17733	Human thymosin beta-4 mRNA, complete cds
180	CATGTTTCCCTCAAA	H1037777	0	1	0	13	1	Examples	N78832	zb17a08.s1 Homo sapiens cDNA clone 302294 3'
									AA411095	zt33d02.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 724131 3'
									W81693	zd84g11.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347396 3'

S1	CATGTTTCCTTCCTT	H1038296	0	6	3	7	17	Examples	M20471	Human brain-type clathrin light-chain a mRNA
									M20472	Human lymphocyte clathrin light-chain A mRNA
1S2	CATGTTTGCACCTTT	H1041504	2	0	0	16	1	Examples	X78947	H.sapiens mRNA for connective tissue growth factor
									U14750	Human connective tissue growth factor mRNA
1S4	CATGTTTGTTAAAA	H1044225							H06492	yl78c08.s1 Homo sapiens cDNA clone 44273 3'
									T35952	EST94173 Homo sapiens cDNA 3' end similar to None
									AA253218	zr53g10.s1 Soares NbHMPu S1 Homo sapiens cDNA clone 667170 3'

Table 5 - Transcripts increased in pancreas and colorectal cancer

SAGE tag that were elevated in both in coloreactal and pancreatic tumor, and are likely to be specific for tumor in general.

	Tag_Sequence		Tag_Number	Accession	Description
1	CATG TGGAAATGAC	C	-950498	M10629	Human alpha-1 collagen gene, 3' end with polyA sit
2	CATG CACTTCAAGG	G	-294155	U42376	Human retinoic acid induced RIG-E precursor (E) mR
				U56145	Human thymic shared antigen-1/stem cell antigen-2
3	CATG ATGTGAAGAG	T(A)	-243747	J03040	Human SPARC/osteonectin mRNA, complete cds.
				M25746	Human osteonectin gene exon 10, complete cds.
4	CATG GCCCAAGGAC	C	-610466	X53416	Human mRNA for actin-binding protein (filamin) (AB
5	CATG ATCTTGTTAC	T	-229106	X02761	Human mRNA for fibronectin (FN precursor).
				K00799	human fibronectin (fn) 3' coding region and flank,
6	CATG GTGCGCTGAG	C	-760291	X58536	Human mRNA for HLA class I locus C heavy chain.
				M26432	Human MHC class I HLA-C.1 gene, complete cds.
7	CATG ACAGGCTACG	G	-76231	M95787	Human 22kDa smooth muscle protein (SM22) mRNA, com
				M83106	Human SM22 mRNA, 5' end.
8	CATG GTGTGTTTGT	A	-769020	M77349	Human transforming growth factor-beta induced gene
9	CATG GATTTCTCAG	C	-589267	X53279	Human mRNA for placental-like alkaline phosphatase
				X55958	H.sapiens mRNA for alkaline phosphatase.
				J04948	Human alkaline phosphatase (ALP-1) mRNA, complete
10	CATG ACCATTCTGC	T	-85882	X57351	Human 1-8D gene from interferon-inducible gene fam
				X02490	Human interferon-inducible mRNA (cDNA 1-8).
11	CATG TCCTTCTCCA	C	-884181	X15804	Human mRNA for alpha-actinin.
12	CATG CTTCTGTGTA	C,T	-515821	D80012	Human mRNA for KIAA0190 protein.
13	CATG ATGTAAAAAA	T	-241665	M74090	Human TB2 gene mRNA, 3' end.
				J03801	Human lysozyme mRNA, complete cds with an Alu repe
				M19045	Human lysozyme mRNA, complete cds.
14	CATG GGCAGAGGAC	C	-673954	X17620	Human mRNA for Nm23 protein, involved in developme
				X75598	H.sapiens nm23H1 gene.
15	CATG AATATTGAGA	A	-53129	U62962	Human Int-6 mRNA, complete cds.
16	CATG TTTTTGATAA	A	-1048113	D16891	Human HepG2 3' region cDNA, clone hmd2cl1.
17	CATG CAGCTGGCCA	T	-302741	X53743	H.sapiens mRNA for fibulin-1 C.

18	CATG GTTCACATTA	G	-774461	X00497	Human mRNA for HLA-DR antigens associated invarian
				M13560	Human Ia-associated invariant gamma-chain gene, ex
19	CATG AAAAGAAACT	T	-2056	Y00345	Human mRNA for polyA binding protein.
20	CATG AATGCAGGCA	G	-58533	M61831	Human S-adenosylhomocysteine hydrolase (AHCY) mRNA
				M61832	Human S-adenosylhomocysteine hydrolase (AHCY) mRNA
21	CATG TGAAATAAAA	C	-918273	X16934	Human hB23 gene for B23 nucleophosmin.
				M28699	Homo sapiens nucleolar phosphoprotein B23 (NPM1) m
				M23613	Human nucleophosmin mRNA, complete cds.
				M26697	Human nucleolar protein (B23) mRNA, complete cds.
22	CATG TTATGGGATC	T	-998030	M24194	Human MHC protein homologous to chicken B complex
23	CATG CAATAAATGT	T	-274492	D23661	Human mRNA for ribosomal protein L37, complete cds
				L11567	Homo sapiens ribosomal protein L37 mRNA, complete
24	CATG AGCCTTTGTT	G	-155632	D83174	Human mRNA for collagen binding protein 2.
25	CATG ACCTGTATCC	C	-97078	X57352	Human 1-8U gene from interferon-inducible gene fam
26	CATG TTCAATAAAA	A	-1000193	M17886	Human acidic ribosomal phosphoprotein P1 mRNA, com
				J05068	human transcobalamin I mRNA, complete cds.
27	CATG CGACCCCACG	C	-398663	M12529	Human apolipoprotein E mRNA, complete cds.
				K00396	Human apolipoprotein E (epsilon 2 and 3 alleles) m
28	CATG CAGATCTTTG	T	-298495	X56998	Human UbA52 adrenal mRNA for ubiquitin-52 amino ac
				X56999	Human UbA52 placental mRNA for ubiquitin-52 amino
29	CATG CTGGCGAGCG	C	-501287	X07491	Human DNA inserts showing sperm-specific hypomethy
				M91670	Human ubiquitin carrier protein (E2-EPF) mRNA, com
30	CATG ATTGGCTTAA	A	-256497	L14272	Human prohibitin (PHB) gene, exons 1-7.
				S85655	prohibitin [human, mRNA, 1043 nt].
31	CATG GTGGTGGACA	C	-765573	U62435	Human nicotinic acetylcholine receptor alpha6 subu
				U68041	Human breast and ovarian cancer susceptibility pro
			-883029	M24398	Human parathymosin mRNA, complete cds.
32	CATG TCCTGCCCCA	T	-125661	X58965	H.sapiens RNA for nm23-H2 gene.
33	CATG ACTGGGTCTA	T		M36981	Human putative NDP kinase (nm23-H2S) mRNA, complet
				L16785	Homo sapiens c-myc transcription factor (puf) mRNA
34	CATG AAGAAGATAG	A	-33331	U02032	Human ribosomal protein L23a mRNA, partial cds.
				U37230	Human ribosomal protein L23a mRNA, complete cds.
				U43701	Human ribosomal protein L23a mRNA, complete cds.

				L13799	Homo sapiens (clone 01) liver expressed protein mR
35	CATG ACATCATCGA	T	-79065	L06505	Human ribosomal protein L12 mRNA, complete cds.
36	CATG CTGTTGGTGA	T	-507577	D14530	Human homolog of yeast ribosomal protein S28, comp
37	CATG ATTATTTTTC	T	-249854	X57959	H.sapiens mRNA for ribosomal protein L7.
				X57958	H.sapiens mRNA for ribosomal protein L7.
				X52967	Human mRNA for ribosomal protein L7.
				L16558	Human ribosomal protein L7 (RPL7) mRNA, complete c
38	CATG GCTTTTAAGG	A	-655115	L06498	Homo sapiens ribosomal protein S20 (RPS20) mRNA, c
39	CATG GGCAAGAAGA	A	-672265	L19527	Homo sapiens ribosomal protein L27 (RPL27) mRNA, c
				L25346	Homo sapiens ribosomal protein L27 (homologue of r
40	CATG CTCTTCGAGA	A	-490889	Y00433	Human mRNA for glutathione peroxidase (EC 1.11.1.9
				Y00483	Human gene for gluthathione peroxidase.
				X13710	H.sapiens unspliced mRNA for glutathione peroxidas
				X13709	Human gpx1 mRNA for gluthatione peroxidase.
				M21304	Human glutathione peroxidase (GPX1) mRNA, complete
41	CATG CTGTTGATTG	C	-507455	X04347	Human liver mRNA fragment DNA binding protein UPI
				U00947	Human clone C4E 3.2 (CAC)n/(GTG)n repeat-containin
42	CATG CTGGGTTAAT	A	-502724	M81757	H.sapiens S19 ribosomal protein mRNA, complete cds
43	CATG ATGGCTGGTA	T	-239533	X17206	Human mRNA for LLRep3.
44	CATG GATGCTGCCA	A	-583573	X59357	Human mRNA for Epstein-Barr virus small RNAs (EBER
				L21756	Homo sapiens acute myeloid leukemia associated pro
				D17652	Human mRNA for HBp15/L22, complete cds.
				S76343	AML1...EAP (translocation breakpoint) [human, chro
45	CATG CCTTCGAGAT	C	-390692	U14970	Human ribosomal protein S5 mRNA, complete cds.
46	CATG CTCCTCACCT	G	-482584	U16811	Human Bak mRNA, complete cds.
				U23765	Human Bak protein mRNA, complete cds.
47	CATG TGTGTTGAGA	G	-978825	X16869	Human mRNA for elongation factor 1-alpha (clone CE
				X16872	Human DNA for elongation factor 1-alpha (clone lam
				X03558	Human mRNA for elongation factor 1 alpha subunit (
				D17182	Human HepG2 3' region MboI cDNA, clone hmd2h03m3.
				D17245	Human HepG2 3' region MboI cDNA, clone hmd4h05m3.
				D17259	Human HepG2 3' region MboI cDNA, clone hmd5d07m3.
				D17276	Human HepG2 3' region MboI cDNA, clone hmd6a12m3.

				M27364	Human elongation factor 1 alpha mRNA, 3' end.
				M29548	Human e'ɔngation factor 1-alpha (EF1A) mRNA, parti
				L41490	Homo sapiens oncogene PTI-1 mRNA, complete cds.
				L41498	Homo sapiens oncogene PTI-1 mRNA, complete cds.
48	CATG TTACCATATC	A	-988366	U57846	Human ribosomal protein L39 mRNA, complete cds.
49	CATG GCCTGCTGGG	C	-621035	X71973	H.sapiens GPx-4 mRNA for phospholipid hydroperoxid
50	CATG CCTCGGAAAA	T	-383489	Z26876	H.sapiens gene for ribosomal protein L38.
51	CATG TACAAGAGGA	A	-803369	X69391	H.sapiens mRNA for ribosomal protein L6.
			-803369	D17554	Human mRNA for DNA-binding protein, TAXREB107, com
			-803369	S71022	neoplasm-related C140 product {human, thyroid carc
52	CATG AACGACCTCG	T	-24951	V00598	Human beta-tubulin pseudogene.
			-24951	V00599	Human mRNA fragment encoding beta-tubulin. (from c
53	CATG CCCTGCCTTG	T	-358783	X55110	Human mRNA for neurite outgrowth-promoting protein
54	CATG CCCAGGGAGA	A	-346761	U38846	Human stimulator of TAR RNA binding (SRB) mRNA, co
				D16933	Human HepG2 3' region cDNA, clone hmd4f11.
55	CATG AGCACCTCCA	G	-148949	Z11692	H.sapiens mRNA for elongation factor 2.
56	CATG CGCCGGAACA	C	-416261	X73974	H.sapiens HRPL4 mRNA.
				D23660	Human mRNA for ribosomal protein, complete cds.
57	CATG CTAAAAAAAA	A	-458753	M33680	Human 26-kDa cell surface protein TAPA-1 mRNA, com
58	CATG GGCTGATGTG	G	-686319	U09510	Human glycyl-tRNA synthetase mRNA, complete cds.
				U09587	Human glycyl-tRNA synthetase mRNA, complete cds.
				D30658	Human T-cell mRNA for glycyl tRNA synthetase, comp
59	CATG ATTCTCCAGT	A	-253260	X55954	Human mRNA for HL23 ribosomal protein homologue.
				X52839	Human mRNA for ribosomal protein L17.
60	CATG GAAAAATGGT	T	-524524	X61156	H.sapiens mRNA for laminin-binding protein.
				X15005	Human mRNA for potential laminin-binding protein (
				U43901	Human 37 kD laminin receptor precursor/p40 ribosom
				J03799	Human colin carcinoma laminin-binding protein mRNA
				M14199	Human laminin receptor (2H5 epitope) mRNA, 5' end.
61	CATG CAGCTCACTG	A	-302367	D87735	Human mRNA for ribosomal protein L14, complete cds
				L10376	Human (clone CTG-B33) mRNA sequence.
				S80520	CAG-isl 7 (trinucleotide repeat-containing sequenc
62	CATG ATAATTCTTT	G	-200576	U14973	Human ribosomal protein S29 mRNA, complete cds.

				L31610	Homo sapiens (clone cori-1c15) S29 ribosomal prote
63	CATG AATCCTGTGG	A	-55227	Z28407	H.sapiens mRNA for ribosomal protein L8.
64	CATG AATAGGTCCA	A	-51925	M64716	Human ribosomal protein S25 mRNA, complete cds.
65	CATG AAAAAAAAAA	A (C, G, T)	-1	X83412	H.sapiens B1 mRNA for mucin.
				Z32564	H.sapiens FRGAMMA mRNA (819bp) for folate receptor
				Z32633	H.sapiens FRGAMMA' mRNA for folate receptor (817bp
				X76180	H.sapiens mRNA for lung amiloride sensitive Na+ ch
				U08470	Human FR-gamma' mRNA, complete cds.
				U08471	Human folate receptor 3 mRNA, complete cds.
				U48697	Human mariner-like element-containing mRNA, clone
				D28532	Human mRNA for renal Na+-dependent phosphate cotra
				M55914	Human c-myc binding protein (MBP-1) mRNA, complete
				L06175	Homo Sapiens P5-1 mRNA, complete cds.
				S73775	calmitine=mitochondrial calcium-binding protein [h
				S77393	transcript ch138 [human, RF1,RF48 stomach cancer c
				X60036	H.sapiens mRNA for mitochondrial phosphate carrier
66	CATG CCAGAACAGA	C	-335945	X79238	H.sapiens mRNA for ribosomal protein L30.
				L16991	Human thymidylate kinase (CDC8) mRNA, complete cds
67	CATG AAGGTGGAGG	A	-44683	X80822	H.sapiens mRNA for ORF.
68	CATG CCTAGCTGGA	T	-379369	X52856	Human cyclophilin-related processed pseudogene.
				X52857	Human cyclophilin-related processed pseudogene.
				X52854	Human cyclophilin-related processed pseudogene.
				X52851	Human cyclophilin gene for cyclophilin (EC 5.2.1.8
				Y00052	Human mRNA for T-cell cyclophilin.
69	CATG GAACACATCC	A	-528694	X63527	H.sapiens mRNA for ribosomal protein L19.
				S56985	ribosomal protein L19 [human, breast cancer cell l
70	CATG AAGGAGATGG	G	-41531	X69181	H.sapiens mRNA for ribosomal protein L31.
				X15940	Human mRNA for ribosomal protein L31.
71	CATG AGGCTACGGA	A	-171113	Z29650	H.sapiens SMCX mRNA.
				D17233	Human HepG2 3' region MboI cDNA, clone hmd4c12m3.
72	CATG AGGTCCTAGC	C	-177610	X08096	Human GST pi gene for glutathione S-transferase pi

			X06547	Human mRNA for class Pi glutathione S-transferase
			X15480	Human mRNA for anionic glutathione-S-transferase (
			X08058	Human glutathione S-transferase pi gene.
			U12472	Human glutathione S-transferase (GST phi) gene, co
			U21689	Human glutathione S-transferase-P1c gene, complete
			U62589	Human glutathione S-transferase P1c (GSTp1c) mRNA,
			M69113	Human fatty acid ethyl ester synthase-III mRNA seq
			M24485	Homo sapiens (clone pHGST-pi) glutathione S-transf
73	CATG TGGTGTTGAG G	-965603	X69150	H.sapiens mRNA for ribosomal protein S18.
			M96153	Homo sapiens apolipoprotein B gene sequence.
			L06432	Homo sapiens 18S ribosomal protein (HKE3) mRNA seq
74	CATG CTCAACATCT C	-475448	M17885	Human acidic ribosomal phosphoprotein P0 mRNA, com
75	CATG GTGTTAACCA G	-769045	L25899	Human ribosomal protein L10 mRNA, complete cds.
76	CATG AGGGCTTCCA A	-174037	X58125	Human (D9S55) DNA segment containing (TG)24 repeat
			D17268	Human HepG2 3' region MboI cDNA, clone hmd5h09m3.
			M73791	Human novel gene mRNA, complete cds.
			M64241	Human Wilm's tumor-related protein (QM) mRNA, comp
			S35960	laminin receptor homolog {3' region} [human, mRNA
77	CATG GGATTTGGCC T	-671654	M17887	Human acidic ribosomal phosphoprotein P2 mRNA, com
			M11147	Human ferritin L chain mRNA, complete cds.
			M12938	Human ferritin light subunit mRNA, partial cds.
			M10119	Human ferritin light subunit mRNA, complete cds.
78	CATG ATTAACAAAG C	-246019	X04409	Human mRNA for coupling protein G(s) alpha-subunit
			X04408	Human mRNA for coupling protein G(s) alpha subunit
			X56009	Human GSA mRNA for alpha subunit of GsGTP binding
			X07036	Human mRNA stimulatory GTP-binding protein alpha s
			M21142	Human guanine nucleotide-binding protein alpha-sub
			M14631	Human guanine nucleotide-binding protein G-s, alph
79	CATG TGTACCTGTA A	-968173	Z36832	H.sapiens (xs31) mRNA, 835bp.
			K00558	human alpha-tubulin mRNA, complete cds.
80	CATG TGGCCCCACC C	-955718	X56494	H.sapiens M gene for M1-type and M2-type pyruvate
			M23725	Human M2-type pyruvate kinase mRNA, complete cds.
			M26252	Human TCB gene encoding cytosolic thyroid hormone-

81	CATG TAATAAAGGT	G	-798764	X67247	H.sapiens rpS8 gene for ribosomal protein S8.
82	CATG GCATAATAGG	T	-602315	X89401	H.sapiens mRNA for large subunit of ribosomal prot
				U14967	Human ribosomal protein L21 mRNA, complete cds.
				U25789	Human ribosomal protein L21 mRNA, complete cds.
				L38826	Homo sapiens L21 ribosomal protein gene, partial c
83	CATG TACCATCAAT	A	-807748	X53778	H.sapiens hng mRNA for uracil DNA glycosylase.
				U34995	Human normal keratinocyte substraction library mRN
				J02642	Human glyceraldehyde 3-phosphate dehydrogenase mRN
				M36164	Human glyceraldehyde-3-phosphate dehydrogenase mRN
				M33197	Human glyceraldehyde-3-phosphate dehydrogenase (GA
84	CATG ATTTGTCCCA	G	-260949	X14957	Human hmgI mRNA for high mobility group protein I.
				X14958	Human hmgI mRNA for high mobility group protein Y.
				M23614	Human HMG-I protein isoform mRNA (HMGI gene), clon
				M23619	Human HMG-I protein isoform mRNA (HMGI gene), clon
				L17131	Human high mobility group protein (HMG-I(Y)) gene
				M23615	Human HMG-Y protein isoform mRNA (HMGI gene), clon
				M23616	Human HMG-Y protein isoform mRNA (HMGI gene), clon
				M23617	Human HMG-Y protein isoform mRNA (HMGI gene), clon
				M23618	Human HMG-Y protein isoform mRNA (HMGI gene), clon
85	CATG GAGGGAGTTT	C	-567488	U14968	Human ribosomal protein L27a mRNA, complete cds.
86	CATG CGCCGCCGGC	T	-416106	U12465	Human ribosomal protein L35 mRNA, complete cds.
87	CATG GTGAAACCCA	ALL	-753749	Z63072	H.sapiens CpG island DNA genomic Msel fragment, cl
88	CATG GTGAAACCCA	ALL	-753749	X16294	Human repetitive DNA containing interspersed repea
89	CATG AAGACAGTGG	C	-33979	X66699	H.sapiens mRNA for ribosomal protein L37a.
				L06499	Homo sapiens ribosomal protein L37a (RPL37A) mRNA,
				L22154	Human ribosomal protein L37a mRNA sequence.
90	CATG CCCCAGCCAG	T	-348755	X55715	Human Hums3 mRNA for 40S ribosomal protein s3.
				U14990	Human XP1PO ribosomal protein S3 (rpS3) mRNA, comp
				U14991	Human XP2NE ribosomal protein S3 (rpS3) mRNA, comp
				U14992	Human IMR-90 ribosomal protein S3 (rpS3) mRNA, com
				S42658	S3 ribosomal protein [human, colon, mRNA, 826 nt].
91	CATG TGGGCAAAGC	C	-959498	X63526	H.sapiens mRNA for protein homologous to elongatio
				Z11531	H.sapiens mRNA for elongation factor-1-gamma.

				M55409	Human pancreatic tumor-related protein mRNA, 3' en
92	CATG TGAGGGAATA	A	-928269	M10036	Human triosephosphate isomerase mRNA, complete cds
93	CATG GACGACACGA	G	-549145	U58682	Human ribosomal protein S28 mRNA, complete cds.
				M58458	Human ribosomal protein S4 (RPS4X) isoform mRNA, c
				M22146	Human scar protein mRNA, complete cds.
94	CATG AACGCGGCCA	A	-26261	Z23063	Homo sapiens macrophage migration inhibitory facto
				L10612	Human glycosylation-inhibiting factor mRNA, comple
				M95775	Homo sapiens macrophage migration inhibitory facto
				L19686	Homo sapiens macrophage migration inhibitory facto
				M25639	Human migration inhibitory factor (MIF) mRNA, comp
95	CATG TGCACGTTTT	C	-935680	X03342	Human mRNA for ribosomal protein L32.
				K03002	Human mRNA from chromosome 15 gene with homology t
96	CATG CACAAACGGT	A	-278636	U57847	Human ribosomal protein S27 mRNA, complete cds.
				L19739	Homo sapiens metallopanstimulin (MPS1) mRNA, compl
97	CATG GGAGTGGACA	T	-667269	L11566	Homo sapiens ribosomal protein L18 (RPL18) mRNA, c
98	CATG GCCGAGGAAG	G	-615043	Z54999	H.sapiens CpG island DNA genomic Msel fragment, cl
				Z57572	H.sapiens CpG island DNA genomic Msel fragment, cl
				Z56073	H.sapiens CpG island DNA genomic Msel fragment, cl
				X53505	Human mRNA for ribosomal protein S12.
99	CATG GGGGAAATCG	C	-696375	M92381	Human thymosin beta 10 mRNA, complete cds.
				M20259	Human thymosin beta-10 mRNA, complete cds.
100	CATG GCAGCCATCC	G	-599350	U14969	Human ribosomal protein L28 mRNA, complete cds.
				D17257	Human HepG2 3' region MboI cDNA, clone hmd5d04m3.
101	CATG TAAGGAGCTG	A	-796831	X77770	H.sapiens RPS26 mRNA.
				X69654	H.sapiens mRNA for ribosomal protein S26.
102	CATG GGCAAGCCCC	A	-672342	U12404	Human Csa-19 mRNA, complete cds.
				X79239	H.sapiens mRNA for ribosomal protein S13.
				L01124	Human ribosomal protein S13 (RPS13) mRNA, complete
103	CATG GTTCCCTGGC	C	-775658	X65923	H.sapiens fau mRNA.
				U02523	Human FAU1P pseudogene, trinucleotide repeat regio
104	CATG CCGTCCAAGG	G	-374027	M60854	Human ribosomal protein S16 mRNA, complete cds.
	CATG TTGGTCCTCT	G	-1027448	Z12962	H.sapiens mRNA for homologue to yeast ribosomal pr
				S64030	L41 ribosomal protein homolog (clone 7B6) [human,

105	CATG CAAACCATCC	A	-263478	X12883	Human mRNA for cytokeratin 18.
				X12876	Human mRNA fragment for cytokeratin 18.
				X12881	Human mRNA for cytokeratin 18.
				M24842	Human keratin 18 (K18) gene, complete cds.
				M26325	Human cytokeratin 18 mRNA, 3' end.
				M26326	Human keratin 18 mRNA, complete cds.
				M26327	Human cytokeratin 18 mRNA, 3' end.
106	CATG AGCTCTCCCT	G	-161624	X53777	Human L23 mRNA for putative ribosomal protein.
107	CATG AGGTCAGGAG	A(T)	-177315	D86979	Human male bone marrow myeloblast mRNA for KIAA022
				X55923	Human DNA for Alu element P1N6.
				X79699	H.sapiens ALU repeat, 230bp.
				X12544	Human mRNA for HLA class II DR-beta (HLA-DR B).
				Z77989	H.sapiens flow-sorted chromosome 6 HindIII fragmen
				U11831	Human clone 2102V-I chromosome 18p telomeric seque
				U12580	Human Alu repeat sequence A3.
				U12582	Human Alu repeat sequence B2.
				U12583	Human Alu repeat sequence D1.
				U14694	Human Alu-Sb2 repeat, clone HALUSB08.
				U14695	Human Alu-Sb2 repeat, clone HALUSB15.
				U14696	Human Alu-Sb2 repeat, clone HALUSB27.
				U14697	Human Alu-Sb2 repeat, clone HUM-11.
				U14698	Human Alu-Sb2 repeat, clone HSB-8P.
				U14699	Human Alu-Sb2 repeat, clone HUM-9.
				U14700	Human Alu-Sb2 repeat, clone HALUSB35.
				U14701	Human Alu-Sb2 repeat, clone HSB-2P.
				U14704	Human Alu-Sb2 repeat, clone HUM-3.
				U14706	Human Alu-Sb2 repeat, clone HUM-10.
				U14707	Human Alu-Sb2 repeat, clone HUM-7.
				J00120	Human (Lawn) c-myc proto-oncogene, complete coding
				L34653	Homo sapiens platelet/endothelial cell adhesion mo
				M37521	Human XV2c gene.
				S61789	NF1=neurofibromatosis type 1 (deletion breakpoint,
				S73483	phosphorylase kinase catalytic subunit PHKG2 homol

				S75201	cholinesterase {Alu element} [human, Insertion Mut
				S75337	{Y Alu polymorphism, YAP, polymorphic subfamily-3}
108	CATG GGGCTGGGGT	C	-695980	Z49148	H.sapiens mRNA for ribosomal protein L29.
				U10248	Human ribosomal protein L29 (humrpl29) mRNA, compl
				U49083	Human cell surface heparin binding protein HIP mRN
				D16992	Human HepG2 partial cDNA, clone hmd2d02m5.
				D16911	Human HepG2 3' region cDNA, clone hmd3b09.
				J03537	Human ribosomal protein S6 mRNA, complete cds.
				M20020	Human ribosomal protein S6 mRNA, complete cds.
109	CATG ACGTTCTCTT	C	-114144		EST
110	CATG TCTCCATACC	C	-906438		EST
111	CATG GACTGCGTGC	C	-555450		EST
112	CATG CTTAATCCTG	A	-508767		EST
113	CATG GGTTGGCAGG	G	-719435		EST
114	CATG GCCCTCTGCC	A	-613862		EST
115	CATG AACAGAAGCA	A	-18469		EST
116	CATG CTGCCGAGCT	C	-497192		EST
117	CATG TTCCTCGGGC	A	-1007018		EST
118	CATG AACTAATACT	A	-28872		EST
119	CATG TAGATAATGG	C	-822331		EST
120	CATG GCCACACCCC	A,C	-607318		EST
121	CATG GAACCCTGGG	A	-529899		EST
122	CATG AACTAAAAAA	A	-28673		EST
123	CATG GAAATGTAAG	A	-528067		EST
124	CATG ACTCCAAAAA	A	-119809		EST
125	CATG GTTCGTGCCA	A	-777109		EST
126	CATG TTACCTCCTT	C	-989024		EST
127	CATG GCACAAGAAG	A	-594051		EST
128	CATG CCCTGGGTTC	T	-359102		EST
129	CATG GCCTGTATGA	G	-621369		EST
130	CATG CCCGTCCGGA	A	-355689		EST
131	CATG AGGAAAGCTG	C	-163999		EST
132	CATG TCAGATCTTT	G	-861056		EST

133	CATG CCAGGAGGAA	T	-338081	EST
134	CATG TCACCCACAC	C	-857362	EST
135	CATG GTGTTGCACA	A	-769605	EST
136	CATG GCCGTGTCCG	C	-618199	EST

Isolation of partial cDNA (3' fragment) by 3' directed PCR reaction

This procedure is a modification of the protocol described in Polyak et al. (1997) Nature 389:300. Briefly, the procedure uses SAGE tags in PCR reaction such that the resultant PCR product contains the SAGE tag of interest as well as additional cDNA, the length of which is defined by the position of the tag with respect to the 3' end of the cDNA. The cDNA product derived from such a transcript driven PCR reaction can be used for many applications.

RNA from a source believed to express the cDNA corresponding to a given tag is first converted to double-stranded cDNA using any standard cDNA protocol. Similar conditions used to generate cDNA for SAGE library construction can be employed except that a modified oligo-dT primer is used to dreive the first strand synthesis. For example, the oligonucleotide of compositon 5'-B-TCC GGC GCG CCG TTT T CC CAG TCA CGA(30)-3', contains a poly-T stretch at the 3' end for hybridization and priming from poly-A tails, an M13 priming site for use in subsequent PCR steps, a 5' Biotin label (B) for capture to strepavidin-coated magnetic beads, and an AscI restriction endonuclease site for releasing the cDNA from the streptavidin-coated magnetic beads. Theoretically, any sufficiently-sized DNA region capable of hybridizing to a PCR primer can be used as well as any other 8 base pair recognizing endonuclease.

cDNA constructed utilizing this or similar modified oligo-dT primer is then processed exactly as described in U.S. Patent No. (insert) up until adapter ligation where only one adapter is ligated to the cDNA pool. After adapter ligation, the cDNA is released from the streptavidin-coated magnetic beads and is then used as a template for cDNA amplification.

Various PCR protocols can be employed using PCR priming sites within the 3' modified oligo-dT primer and the SAGE tag. The SAGE tag-derived PCR primer employed can be of varying length dictated by 5' extension of the tag into the adaptor sequence. cDNA products are now available for a variety of applications.

This technique can be further modified by: (1) altering the length and/or content of the modified oligo-dT primer; (2) ligating adaptors other than that previously employed within the SAGE protocol; (3) performing PCR from template retained on the streptavidin-coated magnetic beads; and (4) priming first strand cDNA synthesis with non-oligo-dT based primers.

Isolation of cDNA using GeneTrapper or modified GeneTrapper Technology

The reagents and manufacturer's instructions for this technology are commercially available from Life Technologies, Inc., Gaithersburg, Maryland. Briefly, a complex population of single-stranded phagemid DNA containing directional cDNA inserts is enriched for the target sequence by hybridization in solution to a biotinylated oligonucleotide probe complementary to the target sequence. The hybrids are captured on streptavidin-coated paramagnetic beads. A magnet retrieves the paramagnetic beads from the solution, leaving nonhybridized single-stranded DNAs behind. Subsequently, the captured single-stranded DNA target is released from the biotinylated oligonucleotide. After release, the cDNA clone is further enriched by using a nonbiotinylated target oligonucleotide to specifically prime conversion of the single-stranded target to double-stranded DNA. Following transformation and plating, typically 20% to 100% of the colonies represent the cDNA clone of interest. To identify the desired cDNA clone, the colonies may be screened by colony hybridization using the 32P-labeled oligonucleotide as described above for solution hybridization, or alternatively by DNA sequencing and alignment of all sequences obtained from numerous clones to determine a consensus sequence.

The genes which are identified herein as being differentially expressed in normal and cancer cells can be used diagnostically and prognostically. Transcription levels in a test sample suspected of being neoplastic can be determined and compared to the levels in normal colon cells. The test sample may be from any tissue suspected of neoplasia, and particularly from either suspected colorectal or suspected pancreatic cancer cells. The control cells for

the purposes of comparison are normal cells, preferably of the same tissue type as the test sample, *e.g.*, colon cells, or pancreatic duct epithelial cells. Upregulation of transcription or downregulation of transcription is therefore diagnostic of the neoplastic state, depending on what gene is used as a test reagent. Similarly, transcription levels can be monitored to assess patent responses to anti-tumor therapies. Transcription levels will also provide prognostic information. For example, the level of transcription in a test sample can be compared to levels found in *bona fide* normal and tumor cells. More extreme deviations from normal expression levels indicate a poorer prognosis.

Transcription levels can be determined according to any means known in the art. These include, without limitation, Northern blots, nuclear run-on assays, *in vitro* transcription assays, primer extension assays, quantitative reverse transcriptase-polymerase chain reactions (RT-PCR), and hybrid filter binding assays. These techniques are well known in the art. See J.C. Alwine, D.J. Kemp, G.R. Stark, *Proc. Natl. Acad. Sci. U.S.A.* 74, 5350 (1977); K. Zinn, D. Di-Maio, T. Maniatis, *Cell* 34, 865 (1983); G. Veres, R.A. Gibbbs, S.E. Scherer, C.T. Caskey, *Science* 237, 415 (1987).

Similarly, upregulated genes and downregulated genes can be detected by measuring expression of their protein products. This can be done by any means known in the art, including but not limited to Western (immuno) blot, enzyme linked immunoadsorbent assay, radioimmunoassay, and enzyme assay. Such techniques are well known in the art. Protein products can be detected in tissue samples of a test patient, using a suspect sample as a test sample, and a matched normal tissue sample from the same tissue type as a control. If normal tissue is not available then a closely related tissue type can be used. Desirably both the samples being compared will be from the same individual. Alternatively, aberrant expression levels of protein products can be detected in body samples, such as blood, serum, feces, urine, sputum. As a control, a normal matched sample can be used from a healthy individual. Aberrant expression levels of transcripts can also be detected in such body samples, particularly in blood and serum.

Probes for use in the assays for transcription levels of particular genes or sets of genes may be RNA or DNA. The probes will be isolated substantially free of other cellular RNAs or DNAs. If the reagent contains one probe then it will comprise at least 50% of the nucleic acids in the reagent composition. If the reagent contains more than one probe, then the proportion will decrease accordingly, so that specific probes will still comprise at least 50% of the nucleic acids in the reagent composition.

Probes can be labeled according to any means known in the art. These may include radioactive labels, fluorescent labels, enzymatic labels, and binding partner labels such as biotin. Means for labeling and detecting probes are well known in the art. Probes comprise at least 10, 11, 12, 15, 20, or 30 contiguous nucleotides of a selected gene.

This invention provides proteins or polypeptides expressed from the polynucleotides of this invention, which is intended to include wild-type and recombinantly produced polypeptides and proteins from procaryotic and eucaryotic host cells, as well as muteins, analogs and fragments thereof. In some embodiments, the term also includes antibodies and anti-idiotypic antibodies.

It is understood that functional equivalents or variants of the wild-type polypeptide or protein also are within the scope of this invention, for example, those having conservative amino acid substitutions. Other analogs include fusion proteins comprising a protein or polypeptide.

The proteins and polypeptides of this invention are obtainable by a number of processes well known to those of skill in the art, which include purification, chemical synthesis and recombinant methods. Full length proteins can be purified from a colon or pancreatic cell or tissue lysate by methods such as immunoprecipitation with antibody, and standard techniques such as gel filtration, ion-exchange, reversed-phase, and affinity chromatography using a fusion protein as shown herein. For such methodology, see for example Deutscher et al. (1999) Guide To Protein Purification: Methods In Enzymology (Vol. 182, Academic Press). Accordingly, this invention also

provides the processes for obtaining these proteins and polypeptides as well as the products obtainable and obtained by these processes.

The proteins and polypeptides also can be obtained by chemical synthesis using a commercially available automated peptide synthesizer such as those manufactured by Perkin Elmer/Applied Biosystems, Inc., Model 430A or 431A, Foster City. The synthesized protein or polypeptide can be precipitated and further purified, for example by high performance liquid chromatography (HPLC). Accordingly, this invention also provides a process for chemically synthesizing the proteins of this invention by providing the sequence of the protein and reagents, such as amino acids and enzymes and linking together the amino acids in the proper orientation and linear sequence.

Alternatively, the proteins and polypeptides can be obtained by well-known recombinant methods as described, for example, in Sambrook et al., (1989), supra, using the host cell and vector systems described above.

Also provided by this application are the polypeptides and proteins described herein conjugated to a detectable agent for use in the diagnostic methods. For example, detectably labeled proteins and polypeptides can be bound to a column and used for the detection and purification of antibodies. They also are useful as immunogens for the production of antibodies as described below. The proteins and fragments of this invention are useful in an in vitro assay system to screen for agents or drugs, which modulate cellular processes.

The proteins of this invention also can be combined with various liquid phase carriers, such as sterile or aqueous solutions, pharmaceutically acceptable carriers, suspensions and emulsions. Examples of non-aqueous solvents include propyl ethylene glycol, polyethylene glycol and vegetable oils. When used to prepare antibodies, the carriers also can include an adjuvant that is useful to non-specifically augment a specific immune response. A skilled artisan can easily determine whether an adjuvant is required and select one. However, for the purpose of illustration only, suitable adjuvants include, but

are not limited to Freund's Complete and Incomplete, mineral salts and polynucleotides.

This invention also provides a pharmaceutical composition comprising any of a protein, analog, mutein, polypeptide fragment, antibody, antibody fragment or anti-idiotipic antibody of this invention, alone or in combination with each other or other agents, and an acceptable carrier. These compositions are useful for various diagnostic and therapeutic methods.

Antibodies can be generated using the proteins encoded by the transcripts identified by the tags disclosed herein. Use of all or portions of the protein as immunogens is routine in the art. Similarly, fusion proteins can be used as immunogens. Antibodies can be affinity purified using the proteins or portions thereof used as immunogens. Similarly, monoclonal antibodies specifically immunoreactive with the protein sequences of the invention can be generated according to techniques which are well known in the art.

Antibodies can be used analytically to quantitate the expression of particular transcripts identified herein as upregulated or downregulated in cancer. In addition, antibodies can be conjugated or non-covalently linked to cytotoxic agents, such as cytotoxins, radionuclides, chemotherapeutic drugs, etc. Such antibodies can be used therapeutically to specifically target cancer cells in which the protein antigens are upregulated. These include the proteins encoded by the transcripts identified by the tags shown in Tables 2, 4, and 5. Means of making such linked cytotoxic antibodies and of administering the same are well known in the art.

Also provided by this invention is an antibody capable of specifically forming a complex with the proteins or polypeptides as described above. The term "antibody" includes polyclonal antibodies and monoclonal antibodies. The antibodies include, but are not limited to mouse, rat, and rabbit or human antibodies.

Laboratory methods for producing polyclonal antibodies and monoclonal antibodies, as well as deducing their corresponding nucleic acid sequences, are known in the art, see Harlow and Lane (1988) supra and

Sambrook et al. (1989) supra. The monoclonal antibodies of this invention can be biologically produced by introducing protein or a fragment thereof into an animal, e.g., a mouse or a rabbit. The antibody producing cells in the animal are isolated and fused with myeloma cells or heteromyeloma cells to produce hybrid cells or hybridomas. Accordingly, the hybridoma cells producing the monoclonal antibodies of this invention also are provided.

Thus, using the protein or fragment thereof, and well known methods, one of skill in the art can produce and screen the hybridoma cells and antibodies of this invention for antibodies having the ability to bind the proteins or polypeptides.

If a monoclonal antibody being tested binds with the protein or polypeptide, then the antibody being tested and the antibodies provided by the hybridomas of this invention are equivalent. It also is possible to determine without undue experimentation, whether an antibody has the same specificity as the monoclonal antibody of this invention by determining whether the antibody being tested prevents a monoclonal antibody of this invention from binding the protein or polypeptide with which the monoclonal antibody is normally reactive. If the antibody being tested competes with the monoclonal antibody of the invention as shown by a decrease in binding by the monoclonal antibody of this invention, then it is likely that the two antibodies bind to the same or a closely related epitope. Alternatively, one can pre-incubate the monoclonal antibody of this invention with a protein with which it is normally reactive, and determine if the monoclonal antibody being tested is inhibited in its ability to bind the antigen. If the monoclonal antibody being tested is inhibited then, in all likelihood, it has the same, or a closely related, epitopic specificity as the monoclonal antibody of this invention.

The term "antibody" also is intended to include antibodies of all isotypes. Particular isotypes of a monoclonal antibody can be prepared either directly by selecting from the initial fusion, or prepared secondarily, from a parental hybridoma secreting a monoclonal antibody of different isotype by using the sib selection technique to isolate class switch variants using the

procedure described in Steplewski et al. (1985) Proc. Natl. Acad. Sci. 82:8653 or Spira et al. (1984) J. Immunol. Methods 74:307.

This invention also provides biological active fragments of the polyclonal and monoclonal antibodies described above. These "antibody fragments" retain some ability to selectively bind with its antigen or immunogen. Such antibody fragments can include, but are not limited to:

(1) Fab,

(2) Fab',

(3) F(ab')2,

(4) Fv, and

(5) SCA

A specific example of "a biologically active antibody fragment" is a CDR region of the antibody. Methods of making these fragments are known in the art, see for example, Harlow and Lane, (1988) supra.

The antibodies of this invention also can be modified to create chimeric antibodies and humanized antibodies (Oi, et al. (1986) BioTechniques 4(3):214). Chimeric antibodies are those in which the various domains of the antibodies' heavy and light chains are coded for by DNA from more than one species.

The isolation of other hybridomas secreting monoclonal antibodies with the specificity of the monoclonal antibodies of the invention can also be accomplished by one of ordinary skill in the art by producing anti-idiotypic antibodies (Herlyn, et al. (1986) Science 232:100). An anti-idiotypic antibody is an antibody which recognizes unique determinants present on the monoclonal antibody produced by the hybridoma of interest.

Idiotypic identity between monoclonal antibodies of two hybridomas demonstrates that the two monoclonal antibodies are the same with respect to their recognition of the same epitopic determinant. Thus, by using antibodies to the epitopic determinants on a monoclonal antibody it is possible to identify other hybridomas expressing monoclonal antibodies of the same epitopic specificity.

It is also possible to use the anti-idiotype technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the mirror image of the epitope bound by the first monoclonal antibody. Thus, in this instance, the anti-idiotypic monoclonal antibody could be used for immunization for production of these antibodies.

As used in this invention, the term "epitope" is meant to include any determinant having specific affinity for the monoclonal antibodies of the invention. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

The antibodies of this invention can be linked to a detectable agent or label. There are many different labels and methods of labeling known to those of ordinary skill in the art.

The antibody-label complex is useful to detect the protein or fragments in a sample, using standard immunochemical techniques such as immunohistochemistry as described by Harlow and Lane (1988) supra. Competitive and non-competitive immunoassays in either a direct or indirect format are examples of such assays, e.g., enzyme linked immunoassay (ELISA) radioimmunoassay (RIA) and the sandwich (immunometric) assay. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

The coupling of antibodies to low molecular weight haptens can increase the sensitivity of the assay. The haptens can then be specifically detected by means of a second reaction. For example, it is common to use haptens such as biotin, which reacts avidin, or dinitropherryl, pyridoxal, and fluorescein, which can react with specific anti-hapten antibodies. See Harlow and Lane (1988) supra.

The monoclonal antibodies of the invention also can be bound to many different carriers. Thus, this invention also provides compositions containing the antibodies and another substance, active or inert. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such, using routine experimentation.

Compositions containing the antibodies, fragments thereof or cell lines which produce the antibodies, are encompassed by this invention. When these compositions are to be used pharmaceutically, they are combined with a pharmaceutically acceptable carrier.

The present invention also provides a screen for various agents which modulate the expression of a gene in a pancreatic or colon cell. To practice the method in vitro, suitable cell cultures or tissue cultures are first provided. The cell can be a cultured cell or a genetically modified cell in which a trancript from SEQ ID NOS:1-732, or their complements, is expressed. Alternatively, the cells can be from a tissue biopsy. The cells are cultured under conditions (temperature, growth or culture medium and gas (CO_2)) and for an appropriate amount of time to attain exponential proliferation without density dependent constraints. It also is desirable to maintain an additional separate cell culture; one which does not receive the agent being tested as a control.

As is apparent to one of skill in the art, suitable cells may be cultured in microtiter plates and several agents may be assayed at the same time by noting genotypic changes, phenotypic changes or cell death.

When the agent is a composition other than a DNA or RNA, the agent may be directly added to the cell culture or added to culture medium for addition. As is apparent to those skilled in the art, an "effective" amount must be added which can be empirically determined. When the agent is a polynucleotide, it may be directly added by use of a gene gun or

electroporation. Alternatively, it may be inserted into the cell using a gene delivery vehicle or vector as described above.

An agent is a potential therapeutic if it alters the expression of gene in the cell. Altered expression can be detected by assaying for altered mRNA expression or protein expression using the probes, primers and antibodies as described herein.

For the purposes of this invention, an "agent" is intended to include, but not be limited to a biological or chemical compound such as a simple or complex organic or inorganic molecule, a peptide, a protein (e.g. antibody) or a polynucleotide (e.g. anti-sense). A vast array of compounds can be synthesized, for example polymers, such as polypeptides and polynucleotides, and synthetic organic compounds based on various core structures, and these are also included in the term "agent". In addition, various natural sources can provide compounds for screening, such as plant or animal extracts, and the like. It should be understood, although not always explicitly stated that the agent is used alone or in combination with another agent, having the same or different biological activity as the agents identified by the inventive screen. The agents and methods also are intended to be combined with other therapies.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

EXAMPLE 1

This example demonstrates the characterization of the general transcription of human colorectal epithelium, colorectal cancers, and pancreatic cancers.

We used the recently developed SAGE (serial analysis of gene expression) method to identify and quantify a total of 303,706 transcripts derived from human colorectal (CR) epithelium, CR cancers or pancreatic cancers (Table 1A) (*3*). These transcripts represented approximately 48,741

different genes (*4*) that ranged in average expression from 1 copy per cell to as many as 5,300 copies per cell (*5*). The number of different transcripts observed in each cell population varied from 14,247 to 20,471. The bulk of the mRNA mass (75%) consisted of transcripts expressed at more than five copies per cell on average (Table 1B). In contrast, the majority (86%) of transcripts were expressed at less than 5 copies per cell, but in aggregate this low abundance class represented only 25% of the mRNA mass. This distribution was consistently observed among the different samples analyzed and was consistent with previous studies of RNA abundance classes based on RNA-DNA reassociation kinetics (Rot curves) . Monte Carlo simulations revealed that our analyses had a 92% probability of detecting a transcript expressed at an average of three copies per cell (*7*).

Table 1 - Summary of SAGE Analysis

A. Overall Summary

	Normal Colon	Colon Tumors	Colon Cell Lines	Pancreatic Tumors	Pancreatic Cell Lines	Total
Total Tags	62,168	60,878	60,373	61,592	58,695	303,706
Unique Genes[1]	14,721	19,690	17,092	20,471	14,247	48,741
GenBank[2]	8,753 (59)	10,490 (53)	10,193 (60)	11,547 (56)	8,922 (63)	26,339 (54)

[1] Indicates the number of different genes represented by the total tags analyzed (4).

[2] Indicates the number of genes that matched an entry in GenBank. The number in parentheses indicates the corresponding percentage of total unique tags.

Table 1 - Summary of SAGE Analysis

B. Summarized by Abundance Classes*

Copies/Cell	Normal Colon	Colon Tumors	Colon Cell Lines	Pancreatic Tumors	Pancreatic Cell Lines	Total
> 500						
Unique Genes	62 (29)	54 (25)	54 (19)	32 (11)	70 (26)	55 (19)
GenBank	59 (95)	52 (96)	53 (98)	32 (100)	70 (100)	54 (98)
> 50 and ≤ 500						
Unique Genes	645 (28)	470 (21)	618 (27)	657 (29)	585 (27)	595 (26)
GenBank	545 (84)	429 (91)	579 (94)	609 (93)	529 (90)	553 (93)
> 5 and ≤ 50						
Unique Genes	4,569 (27)	5,011 (29)	5,733 (34)	6,146 (36)	4,895 (31)	6,209 (30)
GenBank	2,893 (63)	3,204 (64)	3,682 (64)	4,054 (66)	3,168 (65)	4,241 (68)

≤ 5						
Unique Genes	9,445 (16)	14,155 (25)	10,687 (20)	13,636 (24)	8,697 (16)	41,882 (25)
GenBank	5,256 (56)	6,805 (48)	5,879 (55)	6,852 (50)	5,155 (59)	21,491 (51)

*For unique genes, the first number denotes the number of different genes (4) represented in the indicated abundance class. The number in parentheses indicates the mass fraction (X100) of total transcripts represented by the indicated abundance class. For GenBank entries, the first number indicates the number of different genes that matched an entry in GenBank in the indicated abundance class. The number in parentheses indicates the corresponding percentage of total genes.

Many of the SAGE tags appeared to represent previously undescribed transcripts, as only 54% of the tags matched entries in GenBank (Table 1). Twenty percent of these matching transcripts corresponded to characterized mRNA sequence entries in GenBank, whereas 80% matched uncharacterized EST entries. As expected, the likelihood of a tag being present in the databases was related to abundance; GenBank matches were identified for 98% of the transcripts expressed at more than 500 copies per cell but for only 51% of the transcripts expressed at ≤ 5 copies per cell. Because the SAGE data provide a quantitative assay of transcript abundance, unaffected by differences in cloning or PCR efficiency, these data provide an independent and relatively unbiased estimate of the current completeness of publicly available EST databases.

EXAMPLE 2

This example demonstrates a comparison of the expression pattern of normal colon epithelium and primary colon cancers.

Comparison of expression patterns between normal colon epithelium and primary colon cancers revealed that the majority of transcripts were expressed at similar levels (Fig. 1A). However, the expression profiles also revealed 289 transcripts that were expressed at significantly different levels [$P < 0.01$, (8)]. Of these 289, 181 were decreased in colon tumors compared to normal colon (average decrease 10-fold; Fig. 1B; examples in Fig. 2A). Conversely, 108 transcripts were expressed at higher levels in the colon cancers than in normal colon (average increase 13-fold; Fig. 1C; examples in Fig. 2A). Monte Carlo simulations indicated that the analysis would have detected over 95% of those transcripts expressed at a 6-fold or greater level in normal vs. tumor cells or vice versa (*9*). Because relatively stringent criteria were used for defining differences [$P < 0.01$, (*8*)], the number of differences reported above is likely to be an underestimate.

EXAMPLE 3

This example demonstrates the similarities and differences between cancer cell line transcription and transcription of primary cancer tissues. To determine how many of the 289 differences were independent of the cellular microenvironment of cancers *in vivo*, SAGE data from CR cancer cell lines was compared to that from primary CR cancer tissues (Fig. 1B, 1C). Perhaps surprisingly, the majority of transcripts (130 of 181) that were expressed at reduced levels in cancer cells in vivo were also expressed at significantly lower levels in the cell lines (Fig. 1B). Likewise, a significant fraction of the transcripts expressed at increased levels in primary cancers were also expressed at higher levels in the CR cancer cell lines (Fig. 1C). Thus, many of the gene expression differences that distinguish normal from tumor cells in vivo persist during in vitro growth. However, despite these similarities there were also many differences. For example, only 47 of 228 genes expressed at higher levels in CR cancer cell lines were also expressed at high levels in the primary CR cancers.

In combination, comparing the expression pattern of CR cancer cells (in vivo or in vitro) to normal colon revealed 548 differentially expressed transcripts (Fig. 1B,C, Tables 2 and 3). The average difference in expression for these transcripts was 15 fold. Although the ability to detect differences is influenced by the magnitude of the variance with the power to detect smaller differences being less, 92 transcripts that were less than three fold different were identified among the 548 transcripts. However, those genes exhibiting the greatest differences in expression are likely to be the most biologically important.

EXAMPLE 4

This example demonstrates the similarities and differences between colorectal cancer transcription and pancreatic cancer transcription.

To determine whether the changes noted in CR cancers were neoplasia or cell type specific, we performed SAGE on mRNA derived from pancreatic cancers. A total of 404 transcripts were expressed at higher levels in pancreatic cancers compared to normal colon epithelium (examples in Fig. 2B). The majority (268) of these transcripts were pancreas-specific (*10*) (Example in Fig. 2C) although 136 were also expressed at high levels in CR cancers. These 136 transcripts constituted 47% of the 289 transcripts increased in CR cancers relative to normal colon and are likely to be related to the neoplastic process rather than to the specific cell type of origin.

EXAMPLE 5

This example demonstrates the reproducibility of the transcription patterns observed among a larger number of cancer samples.

One question that arose from these data is the potential heterogeneity of expression between individual tumors. The SAGE data were acquired from two examples of each tissue type (normal colon, primary CR cancer, CR cancer cell line, etc.). To examine the generality of these expression profiles, we arbitrarily selected 27 differentially expressed transcripts and evaluated them in six to twelve samples of normal colon and primary cancers by Northern blot analysis (*11*). In general, expression patterns were very reproducible among different samples. Of 10 genes with elevated expression in normal colon relative to CR cancers as determined by SAGE, each was detected in the normal colon samples and was expressed at considerably lower levels in tumors (examples in Fig. 2A). Similarly, most of the genes identified by SAGE as increased in CR or pancreatic cancers were confirmed to be reproducibly expressed in the majority of primary cancers examined by Northern blot (examples in Fig. 2A). It is important to note, however, that there were differences among the cancers, with a few cancers exhibiting particularly high or low levels of individual transcripts. Such differences in gene expression

undoubtedly contribute to the observed heterogeneity in biological properties of cancers derived from the same organ .

EXAMPLE 6

This example demonstrates the identities of some of the transcripts which were found to be differentially expressed in tumor and normal tissues. What are the identities of the differentially expressed genes? Of the 548 differentially expressed transcripts, 337 were tentatively identified through database comparisons. When tested, the great majority (93%) of these identifications proved to be legitimate (*13*), as expected from previous SAGE analyses . Although a large number of differentially expressed genes were identified, some simple patterns did emerge. For example, genes that were expressed at higher levels in normal colon epithelium than in CR tumors were often differentiation-related. These genes included liver fatty acid binding protein , cytokeratin 20 , carbonic anhydrase , guanylin and uroguanylin , which are known to be important for the normal physiology or architecture of the colon epithelium (Table 2). On the other hand, genes that were increased in CR cancers were often related to the robust growth characteristics that these cells exhibit. For example, gene products associated with protein synthesis, including 48 ribosomal proteins, five elongation factors, and five genes involved in glycolysis were observed to be elevated in both CR and pancreatic cancers compared to normal colon cells. Although the majority of the transcripts could not have been predicted to be differentially expressed in cancers, several have previously been shown to be dysregulated in neoplastic cells. The latter included IGFII , B23 nucleophosmin, the Pi form of glutathione S-transferase, and several ribosomal proteins which were all increased in cancer cells as previously reported. Likewise, Dra and gelsolin were both decreased in cancer as previously reported. Surprisingly, two widely studied oncogenes, *c-fos* and *c-erbb3*, were expressed at much higher levels in normal colon epithelium than CR cancers, in contrast to their up-regulation in transformed cells .

In summary, these data provide basic information necessary for understanding the gene expression differences that underlie cancer phenotypes. They additionally provide a necessary framework for interpreting the significance of individual differentially expressed genes. Although this study demonstrated that a large number of such differences exist (approximately 500 at the depth of analysis employed), it was equally remarkable that the fraction of transcripts exhibiting significant differences was relatively small, representing 1.5 % of the transcripts detected in any given cell type (*26*). The fact that many, but not all, of the differences were preserved during in vitro culture demonstrates the utility of cultured lines for examination of some aspects of gene expression, but also provides a note of caution in relying on such lines to perfectly mimic tumors in their natural environment. Finally, the finding that hundreds of specific genes are expressed at different levels in CR cancers, and that some of these are also expressed differentially in pancreatic cancers, provides a wealth of new reagents for future biologic and diagnostic experimentation.

REFERENCES AND NOTES

1. M. D. Adams, *et al.*, *Nature* **377**, **supp. 28**, 3 (1995); M. Schena, D. Shalon, R. W. Davis, P. O. Brown, *Science* **270**, 467 (1995); J. Derisi, *et al.*, *Nature Genetics* 14, 457 (1996); T. M. Gress, *et al.*, *Oncogene* **13**, 1819 (1996); D. J. Lockhart, *et al.*, *Nature Biotechnology* **14**, 1675 (1996); M. Schena, *et al.*, *Proc Natl Acad Sci U S A* **93**, 10614 (1996).

2. V. E. Velculescu, L. Zhang, B. Vogelstein, K. W. Kinzler, *Science* **270**, 484 (1995); V. E. Velculescu, *et al.*, *Cell* **88**, 243 (1997).

3. To minimize individual variation, approximately equal numbers of tags (30,000) were derived from two different patients for each tissue. For primary tumors (two CR carcinomas and two pancreatic adenocarcinomas), RNA was isolated from portions of tumors judged to contain 60%-90% tumor cells by histopathology. The cells grown in vitro were derived from CR (SW837, Caco2) and pancreatic (ASPC-1, PL45) cancer cell lines. CR epithelial cells were isolated from sections of normal colon mucosa from two patients using EDTA as previously described [S. Nakamura, I. Kino, S. Baba, *Gut* 34, 1240 (1993)]. Histopathology confirmed that the isolated cells were greater than 90% epithelial. Isolation of Poly-A RNA and SAGE was performed as previously described (2). SAGE data was analyzed by means of SAGE software and GenBank Release 95 as previously described (2).

4. A total of 69,393 different SAGE tags were identified among the 303,706 tags analyzed. A small fraction of these different tags were likely due to sequencing errors. SAGE analysis of yeast (2), wherein the entire genomic sequence is known, demonstrated a sequencing error rate of ~ 0.7%, translating to a SAGE tag error rate of 6.8% ($1 - 0.993^{10}$). Because these sequencing mistakes are essentially random, they do not substantially affect the analysis although they could artificially inflate the number of unique genes identified. Therefore, to be conservative, we reduced our estimate of unique genes identified by this maximum tag error rate (e.g., 6.8% of 303,706 total tags). The number of different tags derived from the same gene due to alternative splicing was assumed to be negligible.

5. Abundances can be simply determined by dividing the observed number of tags for a given transcript by the total number of tags obtained. An estimate of approximately 300,000 transcripts per cell was used to convert the abundances to copies per cell [N. D. Hastie, J. O. Bishop, *Cell* 9, 761 (1976)].

6. J. O. Bishop, J. G. Morton, M. Rosbash, M. Richardson, *Nature* **250**, 199 (1974); B. Lewin, Gene Expression Vol 2 (John Wiley and sons, New York 1980).

7. Computer simulations indicated that analysis of 300,000 tags would yield a 92 % chance of detecting a tag for a transcript whose expression was at least three copies per cell on average among the tissues examined and assuming 300,000 transcripts per cell.

8. To minimize the number of assumptions and to account for the large number of comparisons being made, Monte Carlo analysis was used for determining statistical significance. The null hypothesis was that the level, kind, and distribution of transcripts were the same for cancer and normal cells. For each transcript, 100,000 simulations were performed to determine the relative likelihood due to chance alone ("p-chance") of obtaining a difference in expression equal to or greater than the observed difference, given the null hypothesis. This likelihood was converted to an absolute probability value by simulating 40 experiments in which a representative number of transcripts (27,993 transcripts in each experiment) was identified and compared. The distribution of transcripts used for these simulations was derived from the average level of expression observed in the original samples. The distribution of the p-chance scores obtained in the 40 simulated experiments (false positives) was then compared to those obtained experimentally. Based on this comparison, a maximum value of 0.0005 was chosen for p-chance. This yielded a false positive rate that was no higher than 0.01 for the least significant p-chance value below the cutoff.

9. Two hundred simulations assuming an abundance of 0.0001 in one sample and 0.0006 in a second sample revealed a significant difference ($P < 0.01$, [8]) 95% of the time.

10. It is not possible to obtain pancreatic ductal epithelium, from which pancreatic carcinomas arise, in sufficient quantities to perform SAGE. It is therefore not possible to determine whether these transcripts were derived from genes that were highly expressed only in pancreatic cancers or were also expressed in pancreatic duct cells.

11. Total RNA isolation and Northern blot analysis was performed as described [W. S. el-Deiry, *et al.*, *Cell* 75, 817 (1993)].

12. A. H. Owens, D. S. Coffey, S. B. Baylin, Eds., Tumor Cell Heterogeneity: Origins and Implications (Academic Press, New York, 1982).

13. Northern blot analyses were done on 45 of the 337 differentially expressed transcripts with tentative database matches. In three cases, the pattern of expression was not differentially expressed as predicted by SAGE and, for the purposes of this calculation, were presumed to represent incorrect database matches.

14. D. C. Rubin, D. E. Ong, J. I. Gordon, *Proc Natl Acad Sci U S A* 86, 1278 (1989); K. Okubo, J. Yoshii, H. Yokouchi, M. Kameyama, K. Matsubara, *DNA Res* 1, 37 (1994).

15. R. Moll, *et al.*, *Differentiation* 53, 75 (1993).

16. J. Sowden, S. Leigh, I. Talbot, J. Delhanty, Y. Edwards, *Differentiation* 53, 67 (1993).

17. F. J. de Sauvage, *et al.*, *Proc Natl Acad Sci U S A* 89, 9089 (1992).

18. R. C. Wiegand, *et al.*, *FEBS Lett* 311, 150 (1992).

19. J. V. Tricoli, *et al.*, *Cancer Res* 46, 6169 (1986); S. Lambert, J. Vivario, J. Boniver, R. Gol-Winkler, *Int J Cancer* 46, 405 (1990).

20. W. Y. Chan, *et al.*, *Biochemistry* 28, 1033 (1989).

21. J. D. Hayes, D. J. Pulford, *Crit Rev Biochem Mol Biol* 30, 445 (1995).

22. G. F. Barnard, *et al.*, *Cancer Res* 52, 3067 (1992); P. J. Chiao, D. M. Shin, P. G. Sacks, W. K. Hong, M. A. Tainsky, *Mol Carcinog* 5, 219 (1992); N. Kondoh, C. W. Schweinfest, K. W. Henderson, T. S. Papas,

Cancer Res 52, 791 (1992); G. F. Barnard, *et al.*, *Cancer Res* 53, 4048 (1993); M. G. Denis, *et al.*, *Int J Cancer* 55, 275 (1993); J. M. Frigerio, *et al.*, *Hum Mol Genet* 4, 37 (1995).

23. C. W. Schweinfest, K. W. Henderson, S. Suster, N. Kondoh, T. S. Papas, *Proc Natl Acad Sci U S A* 90, 4166 (1993).

24. M. Tanaka, *et al.*, *Cancer Res* 55, 3228 (1995); D. Medina, F. S. Kittrell, C. J. Oborn, M. Schwartz, *Cancer Res* 53, 668 (1993).

25. A. D. Miller, T. Curran, I. M. Verma, *Cell* 36, 51 (1984); M. H. Kraus, W. Issing, T. Miki, N. C. Popescu, S. A. Aaronson, *Proc Natl Acad Sci U S A* 86, 9193 (1989).

26. In the case of normal and neoplastic colon cancer tissue, 548 differentially transcripts were identified among the 36,125 unique transcripts.

27. All references cited are hereby incorporated by reference herein.

28. Sequences tags in Tables 2-4 are consecutively numbered to form SEQ ID NOS: 1-732.

CLAIMS

1. A method of diagnosing colon cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 3;

identifying the first sample as neoplastic when the level of the at least one transcript is found to belower in the first sample than in the second sample.

2. A method of diagnosing colon cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

3. The method of claim 1 wherein a comparison of at least two of said transcripts is performed.

4. The method of claim 2 wherein a comparison of at least two of said transcripts is performed.

5. The method of claim 1 wherein a comparison of at least five of said transcripts is performed.

6. The method of claim 2 wherein a comparison of at least five of said transcripts is performed.

7. The method of claim 1 wherein a comparison of at least ten of said transcripts is performed.

8. The method of claim 2 wherein a comparison of at least ten of said transcripts is performed.

9. The method of claim 1 wherein a comparison of at least twenty of said transcripts is performed.

10. The method of claim 2 wherein a comparison of at least twenty of said transcripts is performed.

11. The method of claim 1 wherein a comparison of at least thirty of said transcripts is performed.

12. The method of claim 2 wherein a comparison of at least thirty of said transcripts is performed.

13. An isolated and purified human nucleic acid molecule which comprises a SAGE tag selected from SEQ ID NO:1-732.

14. The nucleic acid molecule of claim 13 which is a cDNA molecule.

15. The nucleic acid molecule of claim 13 wherein the SAGE tag is located at the 3' end of the molecule, adjacent to the 3'-most NlaIII restriction enzyme site.

16. An isolated nucleotide probe comprising at least 10 nucleotides of a human nucleic acid molecule, wherein the human nucleic acid molecule comprises a SAGE tag selected from SEQ ID NO: 1-732.

17. The probe of claim 16 which comprises the selected SAGE tag.

18. A diagnostic reagent for evaluating neoplasia of a colorectal tissue, comprising at least 2 probes according to claim 16.

19. The diagnostic reagent of claim 18 which comprises at least 5 probes according to claim 16.

20. The diagnostic reagent of claim 18 which comprises at least 10 probes according to claim 16.

21. The diagnostic reagent of claim 18 which comprises at least 20 probes according to claim 16.

22. The diagnostic reagent of claim 18 which comprises at least 30 probes according to claim 16.

23. A diagnostic reagent for evaluating neoplasia of a colorectal tissue, comprising at least 2 probes according to claim 17.

24. A method of diagnosing pancreatic cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a pancreatic tissue suspected of being neoplastic and the second sample is of a normal human colon tissue, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

25. A method of diagnosing cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a tissue suspected of being neoplastic and the second sample is of a normal human tissue of the same tissue type, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 5;

identifying the first sample as neoplastic when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

26. A method to aid in the determination of a prognosis for a colon cancer patient, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic colonic tissue and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 3;

determining a poorer prognosis if the level of the at least one transcript is found to be lower in the first sample than in the second sample.

27. A method to aid in determining a prognosis for a patient with colon cancer, comprising the steps of:

comparing the level of at least one transcript in a first tissue sample to a second sample, wherein the first sample is of a colonic cancer tissue and the second sample is of a normal human colonic tissue, and wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

28. A method of diagnosing colon cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 3;

identifying the first sample as neoplastic when the level of expression of the protein is found to be lower in the first sample than in the second sample.

29. A method of diagnosing colon cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a colonic tissue suspected of being neoplastic and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 2;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

30. A method to aid in determining a prognosis of a patient having pancreatic cancer, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic pancreatic tissue and the second sample is of a normal human colon tissue, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

31. A method to aid in providing a prognosis for a cancer patient, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic tissue and the second sample is of a normal human tissue of the same tissue type, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 5;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

32. A method of diagnosing pancreatic cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of expression of at least one protein encoded by a transcript in a first sample of a tissue to a second sample, wherein the first sample is of a pancreatic tissue suspected of being neoplastic and the second sample is of a normal human colon tissue, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

33. A method of diagnosing cancer in a sample suspected of being neoplastic, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a tissue suspected of being neoplastic and the second sample is of a normal human tissue, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 5;

identifying the first sample as neoplastic when expression of the protein is found to be higher in the first sample than in the second sample.

34. A method to aid in the determination of a prognosis for a colon cancer patient, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic colonic tissue and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 3;

determining a poorer prognosis if the level of expression is found to be lower in the first sample than in the second sample.

35. A method to aid in determining a prognosis for a patient with colon cancer, comprising the steps of:

comparing the level of expression of at least one protein in a first tissue sample to a second sample, wherein the first sample is of a colonic cancer tissue and the second sample is of a normal human colonic tissue, and wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 2;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

36. A method to aid in determining a prognosis of a patient having pancreatic cancer, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic pancreatic tissue and the second sample is of a normal human colon tissue, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

37. A method to aid in providing a prognosis for a cancer patient, comprising the steps of:

comparing the level of expression of at least one protein in a first sample of a tissue to a second sample, wherein the first sample is of a neoplastic tissue and the second sample is of a normal human tissue of the same tissue type, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 5;

determining a poorer prognosis if the level of expression is found to be higher in the first sample than in the second sample.

38. A method of treating a cancer cell, comprising the step of:

administering to a cancer cell an antibody which specifically binds to a protein encoded by a transcript identified by a tag selected from the group consisting of those shown in Tables 2, 4, and 5, wherein the antibody is linked to a cytotoxic agent.

39. An antibody linked to a cytotoxic agent, wherein the antibody specifically binds to a protein encoded by a transcript identified by a tag selected from the group consisting of those shown in Tables 2, 4, and 5.

40. A method of detecting colon cancer in a patient, comprising the steps of:

comparing the level of at least one protein in a first body sample to a second body sample, wherein the first sample is a body sample of the patient and the second sample is of a normal human, wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein is found to be higher in the first sample than in the second sample.

41. A method of detecting pancreatic cancer in a patient, comprising the steps of:

comparing the level of at least one protein encoded by a transcript in a first sample of a tissue to a second sample, wherein the first sample is of the patient and the second sample is of a normal human, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein is found to be higher in the first sample than in the second sample.

42. A method of detecting cancer in a patient, comprising the steps of:

comparing the level of at least one protein in a first sample to a second sample, wherein the first sample is of patient and the second sample is of a normal human, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 5, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one protein is found to be higher in the first sample than in the second sample.

43. A method to aid in determining a prognosis for a patient with colon cancer, comprising the steps of:

comparing the level of at least one protein in a first sample to a second sample, wherein the first sample is of a colonic cancer patient and the second sample is of a normal human, wherein the protein is encoded by a transcript identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein is found to be higher in the first sample than in the second sample.

44. A method to aid in determining a prognosis of a patient having pancreatic cancer, comprising the steps of:

comparing the level of at least one protein in a first sample to a second sample, wherein the first sample is of a pancreatic cancer patient and the second sample is of a normal human, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those shown Table 4, wherein said first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein is found to be higher in the first sample than in the second sample.

45. A method to aid in providing a prognosis for a cancer patient, comprising the steps of:

comparing the level of expression of at least one protein in a first sample to a second sample, wherein the first sample is of a cancer patient and the second sample is of a normal human, wherein said protein is encoded by a transcript identified by a tag selected from the group consisting of those

shown Table 5, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one protein is found to be higher in the first sample than in the second sample.

46. A method of detecting colon cancer in a patient, comprising the steps of:

comparing the level of at least one transcript in a first body sample to a second body sample, wherein the first sample is a body sample of the patient and the second sample is of a normal human, wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

47. A method of detecting pancreatic cancer in a patient, comprising the steps of:

comparing the level of at least one transcript in a first sample of a tissue to a second sample, wherein the first sample is of the patient and the second sample is of a normal human, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

48. A method of detecting cancer in a patient, comprising the steps of:

comparing the level of at least one transcript in a first sample to a second sample, wherein the first sample is of patient and the second sample is of a normal human, wherein said transcript is identified by a tag selected

from the group consisting of those shown Table 5, wherein the first and second body sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

identifying neoplasia when the level of the at least one transcript is found to be higher in the first sample than in the second sample.

49. A method to aid in determining a prognosis for a patient with colon cancer, comprising the steps of:

comparing the level of at least one transcript in a first sample to a second sample, wherein the first sample is of a colonic cancer patient and the second sample is of a normal human, wherein the transcript is identified by a tag selected from the group consisting of those shown in Table 2, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

50. A method to aid in determining a prognosis of a patient having pancreatic cancer, comprising the steps of:

comparing the level of at least one transcript in a first sample to a second sample, wherein the first sample is of a pancreatic cancer patient and the second sample is of a normal human, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 4, wherein said first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

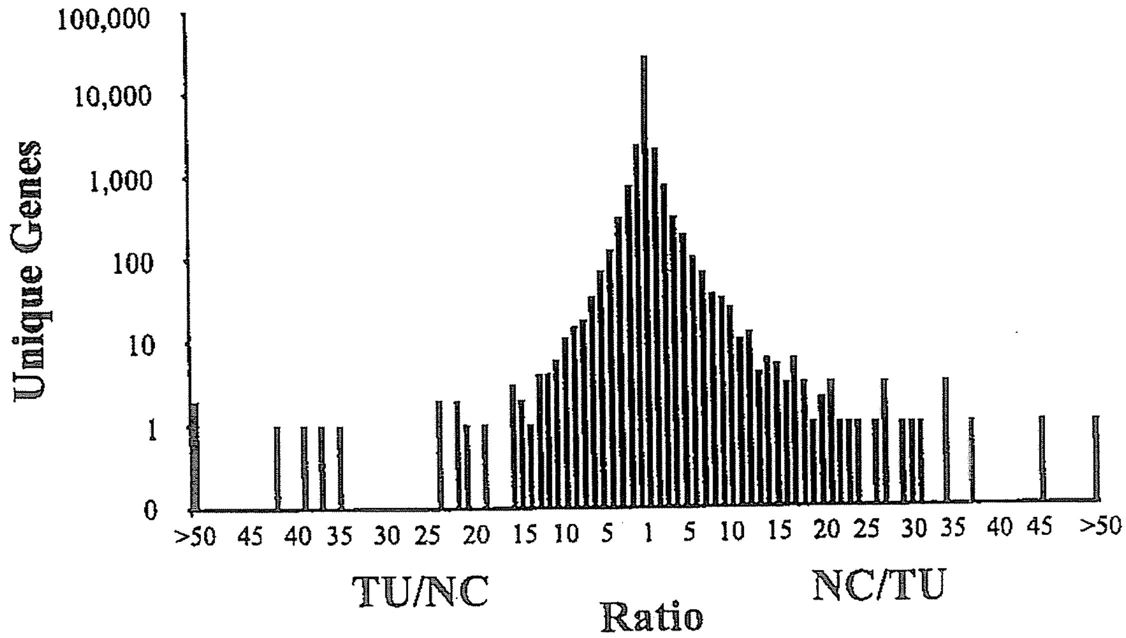
determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

51. A method to aid in providing a prognosis for a cancer patient, comprising the steps of:

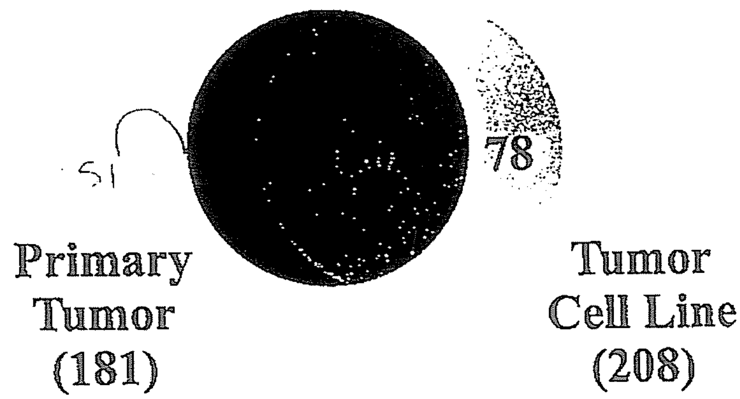
comparing the level of expression of at least one transcript in a first sample to a second sample, wherein the first sample is of a cancer patient and the second sample is of a normal human, wherein said transcript is identified by a tag selected from the group consisting of those shown Table 5, wherein the first and second sample is a sample selected from the group consisting of blood, urine, feces, sputum, and serum;

determining a poorer prognosis if the level of the at least one transcript is found to be higher in the first sample than in the second sample.

52. A method for screening for candidate agents that modulate the expression of a polynuleotide selected from the group consisting of the polynucleotides in SEQ ID NOS:1-732 or their respective complements, comprising contacting a test agent with a colon or pancreatic cell and monitoring expression of the polynucleotide, wherein the test agent which modifies the expression of the polynucleotide is a candidate agent.


100,000
10,000
1,000
100
10
1
0
Unique Genes
>50 45 40 35 30 25 20 15 10 5 1 5 10 15 20 25 30 35 40 45 >50
TU/NC
Ratio
NC/TU

B.


51
78
Primary Tumor (181)
Tumor Cell Line (208)

C.

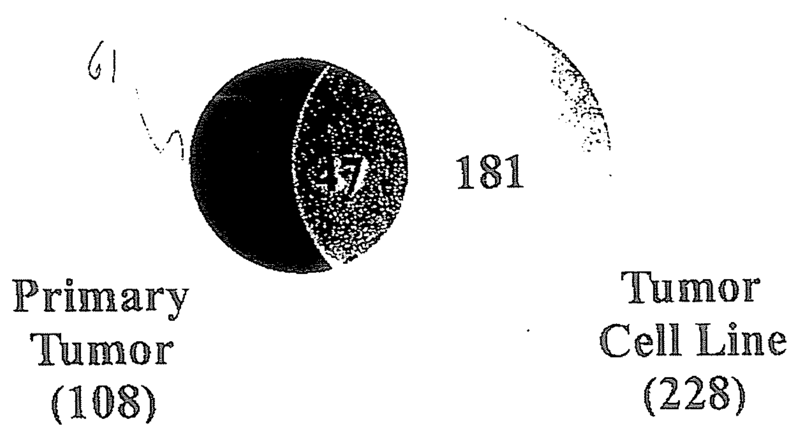
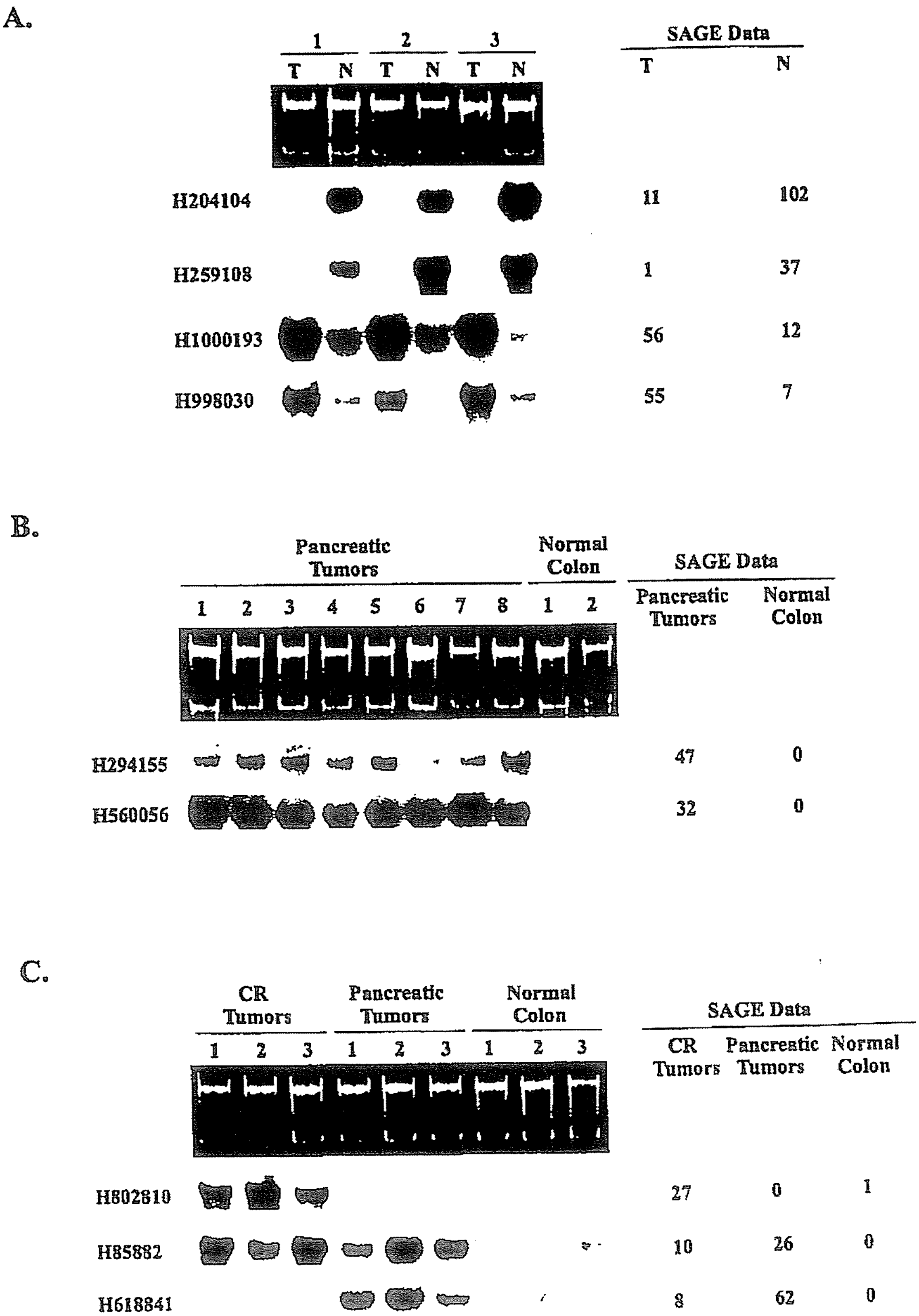

61
47
181
Primary Tumor (108)
Tumor Cell Line (228)

FIG. 2


A.
1
2
3
T N T N T N
SAGE Data
T N
H204104 11 102
H259108 1 37
H1000193 56 12
H998030 55 7
B.
Pancreatic Tumors
Normal Colon
1 2 3 4 5 6 7 8 1 2
SAGE Data
Pancreatic Tumors Normal Colon
H294155 47 0
H560056 32 0
C.
CR Tumors
Pancreatic Tumors
Normal Colon
1 2 3 1 2 3 1 2 3
SAGE Data
CR Tumors Pancreatic Tumors Normal Colon
H802810 27 0 1
H85882 10 26 0
H618841 8 62 0